



جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة كربلاء
كلية الإدارة والاقتصاد
قسم الإحصاء



إستعمال النمذجة الهرمية البيزية للتنبؤ

دراسة تطبيقية عن المصارف

رسالة مقدمة إلى

مجلس كلية الإدارة والاقتصاد في جامعة كربلاء

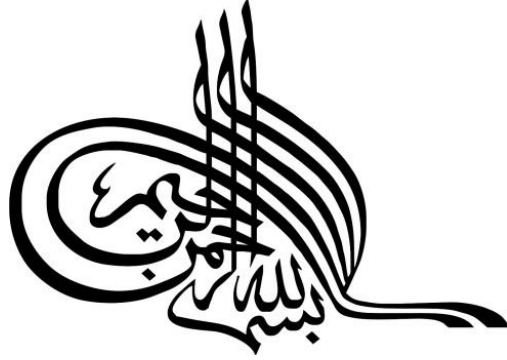
وهي جزء من متطلبات نيل درجة الماجستير في علوم الإحصاء

تقدمت بها

اخلاص عزيز محل

بإشراف

أ.م. د. ايناس عبد الحافظ محمد أ.د. مشتاق كريم عبد الرحيم



﴿وَعِنْدَهُ مَفَاتِحُ الْغَيْبِ لَا يَعْلَمُهَا إِلَّا هُوَ﴾

وَيَعْلَمُ مَا فِي الْبَرِّ وَالْبَحْرِ وَمَا تَسْقُطُ
مِنْ وَرَقَةٍ إِلَّا يَعْلَمُهَا وَلَا حَبَّةٌ فِي
ظُلُمَاتِ الْأَرْضِ وَلَا رَطْبٍ وَلَا يَابِسٍ إِلَّا

فِي كِتَابٍ مُبِينٍ ﴿﴾

صدق الله العلي العظيم

(سورة الانعام: الآية: ٥٩)

إقرار المشرف

أشهد أن إعداد هذه الرسالة الموسومة (استعمال النمذجة الهرمية البيزية للتنبؤ
دراسة تطبيقية عن المصارف) والتي تقدمت بها الطالبة " اخلاص عزيز محل " قد جرت
بإشرافي في قسم الاحصاء / كلية الادارة والاقتصاد / جامعة كربلاء ، وهي جزء من متطلبات نيل
درجة ماجستير علوم في الاحصاء.



أ.م.د. مشتاق كريم عبد الرحيم
بشار خالد
رئيس قسم الاحصاء
2026 / /



أ.م.د. ايناس عبد الحافظ محمد
التاريخ: 2026 / /

توصية رئيس قسم الاحصاء

بناءً على توصية الاستاذ المشرف، أرشح الرسالة للمناقشة.



م.د بشار خالد علي
رئيس قسم الاحصاء
التاريخ: 2026 / /

إقرار الخبير اللغوي

أشهد أن الرسالة الموسومة (استعمال النمذجة الهرمية البيزية للتنبؤ دراسة تطبيقية عن المصارف) قد جرى مراجعتها من الناحية اللغوية تحت اشرافي اذ أصبحت خالية من الاخطاء اللغوية ولأجله وقعت.

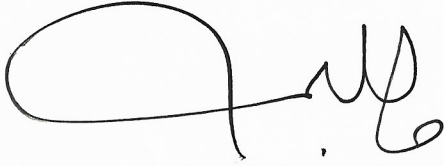
الخبير اللغوي

م.د. صلاح مهدي جابر

2026/ /

إقرار لجنة المناقشة

نشهد نحن أعضاء لجنة المناقشة بأننا قد اطلعنا على الرسالة الموسومة (استعمال النمذجة الهرمية البيزية للتنبؤ دراسة تطبيقية عن المصارف) والمقدمة من قبل الطالبة " اخلاص عزيز محل " وناقشنا الطالبة في محتوياتها وفيما له علاقة بها، ووجدنا بأنها جديرة بنيل درجة ماجستير علوم في الإحصاء بتقدير () .



أ.د. جاسم ناصر حسين

عضواً

2026 / /



أ.د. باسم شلبية مسلم

رئيساً

2026 / /



أ.م.د. ايناس عبد الحافظ محمد

عضواً ومشرفاً

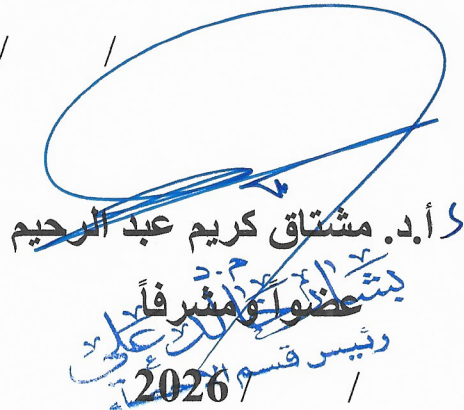
2026 / /



أ.م.د. صدى فايز محمد

عضواً

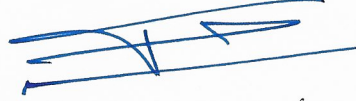
2026 / /



أ.د. مشتاق كريم عبد الرحيم
عضواً ومشرفاً
رئيس قسم المناقشة
2026 / /

إقرار رئيس لجنة الدراسات العليا

بناءً على اقرار المشرف العلمي والخبير اللغوي على رسالة الماجستير للطالبة " اخلص عزيز محل " الموسومة بـ (استعمال النمذجة الهرمية البيزية للتنبؤ دراسة تطبيقية عن المصارف) ارشح هذه الرسالة للمناقشة.



أ.م.د حيدر عباس الجنابي

رئيس لجنة الدراسات العليا

معاون العميد للشؤون العلمية والدراسات العليا

مصادقة مجلس الكلية

صادق مجلس كلية الادارة والاقتصاد/ جامعة كربلاء على قرار لجنة المناقشة.



أ.م.د هاشم جبار الحسيني

عميد كلية الادارة والاقتصاد- جامعة كربلاء

2026 / /

الإهداء ...

بسم الله الرحمن الرحيم

﴿يَرْفَعُ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ﴾

صدق الله العلي العظيم

(المجادلة: الآية: 11)

فالحمد لله حمداً ... اللهم كما أنعمت فزدد وكما زدت فبارك وكما باركت فتمم وكما أتممت فثبتت ...

إلى المبعوث رحمةً للعالمين ذي الخلق العظيم الرسول الأعظم محمد (صلى الله عليه وسلم) واله الطيبين الطاهرين.

إلى أمام زماني المنتظر (عجل الله فرجه) ...

إلى من كان سندي في الحياة، من رحل وبقيت روحه الطاهرة ترعاني.. روح والدي الحبيب الذي زرع في قلبي حب العلم وغرس في نفسي معنى القوة والصبر.. رحمك الله رحمة واسعة

إلى من تغمرني بحنانها وحبها وعطائها، إلى من كانت دعواتها تنقذني في كل موطن وموقف
امي الحبيبة

إلى رفيق دربي وسندي والدا عم الأول الذي كان عوناً وسنداً في كل خطوة، إلى من يلهمني الشجاعة والقوة ويصبرني في المحن والشدائد ويرسم البسمة على وجهي وتبتهج روعي بقربه

إلى (زوجي الحبيب أبو طيبة) شكراً لصبرك وتشجيعك الدائم.

إلى اخوتي وأخواتي الاعزاء. إلى اختي الغالية ام هشام على وجه الخصوص لما كان لها من دعم لا يُفدّر بثمن

إلى صديقاتي الغاليات، مع خالص الامتنان لصديقتي واختي الغالية (فتها م).

إلى اساتذتي الكرام الدكتور مشتاق كريم عبد الرحيم والدكتورة ايناس عبد الحافظ محمد كل الامتنان والعرفان لما منحتماني من علم وتوجيه ودعم

إلى قسمي العزيز قسم الإحصاء حاضن المعرفة وبوابة الطموح لكم كل الحب والتقدير إدارة وتدرسين وزملاء.

لكم جميعاً اهدي ثمرة هذا الجهد المتواضع تقديراً ووفاءً ومحبةً خالصة...

الباحث

شكر وامتنان ...

الحمد لله الذي بنعمته تتم الصالحات، والصلاة والسلام على رسول الله وخاتم النبيين وعلى آله الطيبين الطاهرين وبعد:

أرفع يديّ شكراً و عرفاناً لربِّ كريمٍ لطيفٍ أكرمني بفيض عطاياه، ووهبني الصبر والقوة فله الحمد كما ينبغي لجلال وجهه وعظيم سلطانه.

ثم أنني أقفُ اعترافاً مني بالجميل وأتوجه بجزيل شكري وخالص أمتناني لأصحاب الفضل الكبير أساتذتي الدكتور (ايناس عبد الحافظ محمد) والدكتور (مشتاق كريم عبد الرحيم) الذين تفضلا بالإشراف على رسالتي وأرشادي وتوجيهي، فلهم كل الاعتزاز والاحترام والتقدير وأسأل الله العلي القدير أن يحفظهم وينعم عليهم بالصحة والعافية والتوفيق إنه سميعٌ عليم.

كما أتقدم بالشكر والامتنان إلى الأساتذة الفضلاء في قسم الإحصاء- كلية الإدارة والاقتصاد – جامعة كربلاء لما بذلوه من جهد وعناء في رحلتي الأكاديمية ولدعمهم العلمي المستمر.

كما أتقدم بالشكر الجزيل إلى السادة رؤيس وأعضاء لجنة المناقشة المحترمين لتفضالهم بقبول مناقشة الرسالة ولما سيبدون من توجيهات ونصائح وملاحظات قيمة.

وكما أتقدم بالشكر والامتنان لعائلتي الكريمة، وامتناني لرفيق روحي وسر سعادتي سندي العظيم الذي لا يميل فلهم مني كل الحب والتقدير.

وفي الختام أتقدم بالشكر إلى كل من دعمني وتمنى لي الخير.

ومن الله التوفيق ...

الباحث ...

قائمة المحتويات

الصفحة	الموضوع	
ب	الآية القرآنية	
ت	الإهداء	
ث	شكر وامتنان	
ج-د	قائمة المحتويات	
ذ-ش	قائمة الجداول	
ص-ض	قائمة الاشكال	
ظ	قائمة الرموز والمصطلحات	
ع	المستخلص	
6-1	(منهجية الرسالة والاستعراض المرجعي)	الفصل الأول
1	المقدمة (Introduction)	1.1
2	مشكلة الرسالة (Problem of the thesis)	1.2
2	هدف الرسالة (Aim of the thesis)	1.3
6-2	الاستعراض المرجعي: (Literature Review)	1.4
33-7	(الجانب النظري)	الفصل الثاني
7	تمهيد (Preface)	2.1
7	الأسس النظرية للانحدار التقليدي	2.2
7	الانحدار الخطي (Linear Regression)	2.2.1
8	الانحدار غير الخطي (Non-Linear Regression)	2.2.2
11-8	الانحدار اللوجستي الثنائي المتعدد (Multiple Binary Logistic Regression)	2.2.3

12-11	النماذج متعددة المستويات والبيانات الهرمية (<i>Multilevel Models and Hierarchical Data</i>)	2.3
16-13	انموذج الانحدار اللوجستي (بمستويين) (<i>Two – level Logistic Regression Model</i>)	2.4
16	طرائق التقدير التقليدية <i>Traditional Estimation Methods</i>	2.5
22-16	طريقة الإمكان الأعظم <i>Maximum Likelihood Method (MLE)</i>	2.5.1
23-22	خطوات خوارزمية نيوتن -رافسون لتقدير معلمات الانموذج بطريقة الإمكان الأعظم (MLE)	2.5.2
28-23	الأسلوب البيزي بطريقة (MAP) <i>The Bayesian method in a MAP way</i>	2.6
28	توظيف مقدر الانتروبي مع الأسلوب البيزي بطريقة (MAP)	2.7
28	مفهوم الانتروبي العظمى (<i>The Concept of Maximum Entropy</i>)	2.7.1
31-29	مقدر الانتروبي العظمى للمعلمات <i>Maximum Entropy Estimator for the Parameters</i>	2.7.2
32	خطوات خوارزمية نيوتن رافسون لتقدير المعلمات بطريقة توظيف مقدر الانتروبي في التقدير البيزي بطريقة MAP	2.7.3
33-32	معايير تقييم طرائق التقدير في النماذج اللوجستية	2.8
81-34	(الجانب التجريبي)	الفصل الثالث
34	تمهيد (Preface)	3.1

35-34	مفهوم المحاكاة (Simulation Concept)	3.2
35	مراحل وخطوات المحاكاة (Simulation stages and steps)	3.3
36-35	تحديد القيم الافتراضية لمعاملات النماذج المدروسة في تجارب المحاكاة	3.3.1
37-36	توليد التأثيرات الثابتة (Fixed effects)	3.3.2
37	اختيار تباينات التأثيرات العشوائية (Random effects)	3.3.3
39-37	توليد بيانات المتغيرات (Variables generation)	3.3.4
39	مقارنة النتائج	3.3.5
81-39	تحليل تجارب المحاكاة	3.3.6
95-82	(الجانب التطبيقي)	الفصل الرابع
82	تمهيد (Preface)	4-1
82	البيانات التطبيقية (Applied data)	4-2
83	اختبار تحقق افتراضات نموذج الانحدار اللوجستي متعدد المستويات	4-3
83	التأكد من نوع المتغيرات	4-3-1
84-83	اختبار توزيع المتغيرات	4-3-2
84	اختبار التعدد الخطي بين متغيرات المستوى الاول (Multicollinearity)	4-3-3
86-85	اختبار التعددية بين المستويات (Multilevel Structure)	4-3-4
86	التأكد من خطية الانموذج (Linearity)	4-3-5
87-86	اختبار وجود التفاعل بين المتغيرات (Interactions)	4-3-6
88-87	التحقق من الاستقلالية (Independence)	4-3-7
94-88	تقدير انموذج الانحدار اللوجستي الثنائي بمستويين (Model estimation)	4-4

95-94	التنبؤ بمتغير الاستجابة <i>Prediction of the response variable</i>	4-5
97-96	(الاستنتاجات والتوصيات)	الفصل الخامس
96	الاستنتاجات: (Conclusions)	5-1
97	التوصيات: (Recommendations)	5-2
102-98	المصادر	
98	المصادر العربية (<i>Arabic references</i>)	أولاً:
98 102-	المصادر الأجنبية (<i>Foreign references</i>)	ثانياً:

قائمة الجداول

الصفحة	عنوان الجدول	رقم الجدول
41	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الأول عند حجم عينة $n=25$	(1-3)
42	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الأول عند حجم عينة $n=25$	(2-3)
44	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الأول عند حجم عينة $n=50$	(3-3)
45	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الأول عند حجم عينة $n=50$	(4-3)
47	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الأول عند حجم عينة $n=75$	(5-3)
48	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الأول عند حجم عينة $n=75$	(6-3)
48	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الأول عند حجم عينة $n=100$	(7-3)
50	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الأول عند حجم عينة $n=100$	(8-3)

53-52	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه $n=25$	(9-3)
54	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه $n=25$	(10-3)
57-56	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه $n=50$	(11-3)
58	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه $n=50$	(12-3)
61-60	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه $n=75$	(13-3)
62	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه $n=75$	(14-3)
65-64	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه $n=100$	(15-3)
66	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه $n=100$	(16-3)
69-68	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه $n=25$	(17-3)

70	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه n=25	(18-3)
72	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه n=50	(19-3)
73	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه n=50	(20-3)
75	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه n=75	(21-3)
76	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه n=75	(22-3)
78	المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه n=100	(23-3)
79	الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه n=100	(24-3)
81	عدد مرات الأفضلية حسب الرتبة لكل طريقة حسب احجام العينات وحسب المجموعات وحسب عدد المتغيرات المستقلة في المستوى الاول والثاني.	(25-3)
83	اختبار الاعتدالية لمتغيرات المستوى الاول	1-4
84	عامل تضم التباين لمتغيرات المستوى الاول	2-4

85	نتائج اختبار نسبة الأرجحية بين انموذجين متداخلين (Nested Models)	3-4
86	نتائج اختبار خطية العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة	4-4
87	نتائج اختبار نسبة الأرجحية لوجود التفاعل بين المتغيرات	5-4
87	اختبار الاستقلالية داخل الفروع	6-4
88	المعاملات المقدرة ومستوى الدلالة لكل معلمة من معاملات انموذج الانحدار اللوجستي بمستويين المستوى الاول (العمر والدخل) والمستوى الثاني (رقم الفرع 1, 2, 3 وحجم الفرع (عدد الموظفين)) بموجب طريقة BME عند البيانات الحقيقية	7-4
93-90	القيم الاحتمالية التقديرية لمعادلة انموذج الانحدار اللوجستي متعدد المستويات بمستويين المستوى الاول (العمر والدخل) والمستوى الثاني (رقم الفرع 1, 2, 3 وحجم الفرع (عدد الموظفين)) بموجب طريقة MLE وطريقة BME وطريقة BM عند البيانات الحقيقية ومتوسط مربعات الخطأ ومعايير الدقة	8-4
94	Hosmer-Lemeshow لكل فرع	9-4
94	القيم الاحتمالية للتنبؤ بالسداد والتعثر	10-4

قائمة الأشكال

الصفحة	عنوان الشكل	رقم الشكل
10	تحويل الانموذج من الصيغة الغير خطية الى الصيغة الخطية	(1-2)
13	انموذج متعدد المستويات بمستويين	(2-2)
24	مراحل دالة الكثافة الاحتمالية اللاحقة	(3-2)
43	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=25$.	(1-3)
46	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=50$.	(2-3)
48	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=75$.	(3-3)
51	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=100$.	(4-3)
55	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينه $n=25$.	(5-3)
59	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينه $n=50$.	(6-3)
63	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينه $n=75$.	(7-3)
67	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينه $n=100$.	(8-3)

71	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينه n=25	(9-3)
74	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينه n=50	(10-3)
77	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينه n=75	(11-3)
80	الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينه n=100	(12-3)
93	منحنيات القيم الاحتمالية المقدرة عند طرائق التقدير المختلفة عند البيانات الحقيقية	1-4

قائمة الرموز والمختصرات

الرمز	المعنى
$H(\beta_r)^{-1}$	معكوس مصفوفة هيسين <i>Inverse of the Hessian matrix</i>
$H(\beta_r)$	مصفوفة هيسين (المشتقة الثانية) Hessian matrix (second derivative)
$U(\beta_r)$	المشتقات الجزئية الاولى للمعلمات The first partial derivative of the parameters
BM	Bayes-MAP
BME	Bayes-MAP Entropy
$\pi(\beta)$	التوزيع الاولي (<i>Prior Distribution</i>)
Σ^{-1}	معكوس مصفوفة التباين
$\pi(\beta x)$	التوزيع اللاحق (<i>Posterior Distribution</i>)
∇	متجه المشتقات الجزئية
$\nabla_{\beta} \log \pi(\beta^r x)$	موجه المشتقات الجزئية للوغاريتم التوزيع اللاحق
d_k	موجه نقاط الدعم <i>Support Points Vector</i>
p_k	موجه الاحتمالات المقابلة لنقاط الدعم
\hat{P}_{qk}	الاحتمالات التقديرية المقابلة لنقاط الدعم
GME	مقدر الانتروبي العظمى العامة
H	دالة الهدف (<i>Objective function</i>)
D	مصفوفة نقاط الدعم (<i>Support points matrix</i>)
p	مصفوفة الاحتمالات (<i>Probability Matrix</i>)
l	معادلة لاكرانج (<i>Lagrange equation</i>)

عدد نقاط الدعم (<i>Number of support points</i>)	q
متوسط مربعات الخطأ التكاملي <i>Integrated Mean Squared Error – IMSE</i>	IMSE
الانموذج الكامل الذي يتضمن جميع المتغيرات التوضيحية	LL_{full}
الانموذج الفارغ الذي لا يحتوي على أي متغيرات	L_{null}
توزيع كاي سكوير – <i>Chi</i> <i>Square Distribution</i>	(χ^2)
إحصاءة نسبة الإمكان <i>Likelihood Ratio Statistic</i>	(D)

المستخلص:

تتناول هذه الرسالة نمذجة وتحليل البيانات ذات البنية الهرمية متعددة المستويات باستعمال انموذج الانحدار اللوجستي الثنائي بمستويين ضمن الإطارين التقليدي والبيزي، مع التركيز على تحسين دقة تقدير المعلمات وكفاءة التنبؤ. تكتسب هذه الرسالة أهميتها من الطبيعة الهرمية للبيانات المستخدمة في كثير من التطبيقات الواقعية، ولا سيما في القطاع المصرفي، حيث تتداخل المشاهدات ضمن مجموعات مختلفة، الأمر الذي يجعل الطرائق التقليدية القائمة على مستوى واحد غير كافية لتمثيل مصادر التباين المتعددة بصورة دقيقة. هدفت الرسالة إلى تطوير أسلوب إحصائي متكامل يجمع بين النمذجة الهرمية البيزية وطريقة التقدير البيزي بطريقة ماب (MAP)، مع توظيف مقدر الانتروبي العظمى كقيمة أولية، لتحسين دقة التنبؤات واتخاذ القرارات بشكل أكثر كفاءة. وتمت مقارنة هذه الطريقة المقترحة مع طريقتي الإمكان الأعظم (MLE) وطريقة التقدير البيزي بطريقة ماب BM من حيث الأداء الإحصائي والكفاءة التقديرية.

اعتمد الجانب النظري على استعراض مفصل لأنموذج الانحدار اللوجستي متعدد المستويات، وطرائق تقدير معالمه التقليدية والبيزية، مع بيان دور التوزيعات الأولية ونظرية بيز ومفهوم الانتروبي العظمى وآلية دمجها ضمن BM. أما الجانب التجريبي، فقد تضمن إجراء سلسلة من تجارب المحاكاة باستخدام أسلوب مونت كارلو، شملت سيناريوهات مختلفة لأحجام العينات، وعدد المجموعات، وعدد المتغيرات التوضيحية في المستويين، وتم تقييم أداء الطرائق المعتمدة باستعمال معايير إحصائية مناسبة، أبرزها متوسط مربعات الخطأ التكاملي ومعامل التحديد الزائف لماكفادن واختبار نسبة الإمكان. كما تم تطبيق الانموذج المقترح على بيانات حقيقية جُمعت من ثلاثة فروع لمصرف الرافدين في محافظة كربلاء، لغرض تحليل سلوك سداد القروض والتنبؤ بحالات التعثر المالي، مع الأخذ بعين الاعتبار التأثيرات الفردية وتأثيرات الفروع المصرفية. أظهرت النتائج التجريبية والتطبيقية تفوق طريقة BME من حيث دقة التقدير واستقرار النتائج مقارنة بالطرائق الأخرى، ولا سيما في حالات العينات الصغيرة والبنى الهرمية المعقدة.

الفصل الاول

منهجية الرسالة
والاستعراض المرجعي

1.1 المقدمة (Introduction)

تعد عملية التنبؤ من العمليات المهمة التي يقوم بها الباحث لغرض الاطلاع على الظاهرة المدروسة في المستقبل وتعتمد هذه العملية على نوعية البيانات وطبيعتها فضلاً عن الاساليب الإحصائية التي تستعمل في عملية التنبؤ وهناك بيانات تقليدية تكون على مستوى واحد Level 1 لكل المتغيرات، أي تكون بيانات غير هرمية (non-hierarchical data) فإن اجراء التنبؤ لمثل هكذا بيانات يتم باستعمال نماذج الانحدار إذ يجب الحصول على المعادلة التقديرية عن طريق تقدير معلمات نموذج الانحدار بإحدى الطرائق، ولكن هناك بيانات بمستويات متعددة أي تكون على أكثر من مستوى واحد كمستويين او اكثر وتسمى البيانات الهرمية hierarchical data وتستعمل النمذجة لهذه البيانات اما بطرائق تقليدية Classical hierarchical data والتي تعتمد أحيانا على التقدير بطريقة الإمكان الأعظم (Maximum Likelihood Estimation – MLE) او المربعات الصغرى لكن قد تكون هذه النمذجة محدودة ولا تعطي تنبؤا واضحا عن الظاهرة المدروسة فضلا عن توفر معلومات مسبقية عن معلمات نموذج الانحدار التي من الممكن ان تخدم عملية التنبؤ وبهذا تكون النمذجة بالطرائق التقليدية غير جيدة في التنبؤ؛ لأنها تتسم بإهمالها للمعلومات المسبقة عن المعلمات في عملية التقدير. لذلك يتم اللجوء إلى النمذجة البيزية التي تتيح دمج المعلومات المسبقة في عملية التقدير سعيا لتحسين دقة التنبؤات واتخاذ القرارات بشكل أكثر كفاءة.

لذلك تضمنت الرسالة خمسة فصول: الفصل الأول يشمل المقدمة ومشكلة وهدف الرسالة مع الاستعراض المرجعي الذي تضمن بعض البحوث والدراسات السابقة التي لها علاقة بموضوع الرسالة. اما الفصل الثاني المتمثل بالجانب النظري الذي تضمن اهم الأسس النظرية والمفاهيم الأساسية ذات الصلة بموضوع الرسالة. مع التركيز على طرائق تقدير معلماته المتمثلة بطريقة الإمكان الأعظم MLE باستعمال خوارزمية نيوتن - رافسون وطريقة الأسلوب البيزي بطريقة MAP وكيفية توظيف مقدر الانتروبي كقيمة اولية مع أسلوب بيز MAP. في حين تضمن الفصل الثالث تقديم نتائج عدد من تجارب المحاكاة والتي بلغ عددها (12) تجربة مختلفة وجدول يبين عدد مرات الأفضلية لكل طريقة بحسب احجام العينات وبحسب المجموعات وعدد المتغيرات المستقلة في المستوى الأول والثاني. اما الفصل الرابع تضمن الجانب التطبيقي إذ استعملت بيانات حقيقية تم جمعها من ثلاثة فروع تابعة لمصرف الرافدين في محافظة كربلاء المقدسة. واخيرا الفصل الخامس شمل اهم الاستنتاجات والتوصيات التي توصلت اليها الرسالة.

1.2 مشكلة الرسالة (Problem of the thesis)

نظرًا لطبيعة البيانات المتاحة للدراسة، والتي تتضمن مشاهدات متداخلة ومتكررة ضمن ثلاثة فروع مختلفة، فإن هذه البيانات تتسم ببنية هرمية متعددة المستويات. يُلزم هذا التركيب الهرمي استعمال نماذج إحصائية متخصصة تأخذ في الاعتبار مصادر التباين المختلفة لضمان تقديرات دقيقة وتنبؤات موثوقة لحالات التعثر المالي للعملاء.

وعلى الرغم من شيوع استعمال أساليب التقدير التقليدية إلا أن تجاهلها للمعلومات المسبقة عن معالم الانموذج، رغم توفرها، يؤدي إلى فقدان جزء مهم من المعرفة الإحصائية التي يمكن أن تسهم في تحسين جودة التقدير. وبناءً على ذلك تتمثل المشكلة الأساسية لهذه الرسالة في كيفية نمذجة البيانات الهرمية متعددة المستويات مع الافادة القصوى من المعلومات المسبقة المتاحة والتي تؤدي الى تعزيز دقة التنبؤ بحالات التعثر المالي في القطاع المصرفي.

1.3 هدف الرسالة (Aim of the thesis)

1- توظيف أسلوب احصائي يعالج البيانات الهرمية ويأخذ بعين الاعتبار المعلومات المسبقة لمعلمات الانموذج والذي يتمثل بتوظيف مقدر الانتروبي كقيمة اولية مع الأسلوب البيزي بطريقة MAP (BME) لتقدير معلمات انموذج الانحدار اللوجستي الثنائي بمستويين. ومقارنتها مع طرائق التقدير الأخرى مثل طريقة MLE وطريقة BM لتحديد الطريقة الأمثل لتقدير المعلمات وتطبيقها على البيانات الحقيقية.

2-التنبؤ بحالات التعثر المالي للعملاء في الفروع الثلاثة ، بهدف تحسين دقة التنبؤات عن سداد القروض بناءً على عوامل متعددة.

1.4 الاستعراض المرجعي (Literature Review)

تم عرض أهم الدراسات السابقة المرتبطة بموضوع الرسالة، سواء كانت التي تناولت الجوانب النظرية أم طرائق التقدير الخاصة بالنمذجة الهرمية البيزية. ويُعدّ *Smith* و *Lindley* أول من ناقش هذا النوع من النماذج عام (1972)، ما فتح المجال أمام العديد من الأبحاث اللاحقة. ومنذ عام 1995، ازداد الاهتمام بالنمذجة الهرمية البيزية بشكل واضح، وتطورت طرائق التقدير

المستعملة فيها بفضل التقدم في الأساليب الحاسوبية. وصولاً إلى أحدث التطورات المعاصرة.

ففي عام (1995)، يشير الباحث (Stangl) في اطروحته إلى البدايات التأسيسية لهذا النوع من النمذجة، إذ قدم تطبيقاً رائداً في مجال البيانات السريرية متعددة المراكز. تميز هذا العمل بتقديمه منهجية متطورة لمعالجة التباين بين المجموعات الفرعية عن طريق استعمال النماذج الهرمية البيزية. وقد اعتمد الباحث على افتراض أن المتغير المعتمد من النوع المستمر ولا يتبع التوزيع الطبيعي، مما مكنه من تحليل بيانات المجاميع الفرعية وإيجاد مقدرات البقاء على قيد الحياة. إلا أن هذا العمل بقي محصوراً بالبيانات المستمرة غير الطبيعية، ولم يتناول تعقيدات البنى الهرمية الأكثر تطوراً، ما أشار إلى ضرورة تطوير نماذج أكثر شمولية قادرة على معالجة أنواع البيانات المختلفة والتعقيدات البنيوية الأكثر تطوراً. [33]

في عام 1998، وسع الباحث (Browne) في اطروحته نطاق التطبيق بشكل كبير ليشمل تحليلات معقدة للبيانات متعددة المستويات، فقد ركّز على استغلال تقنيات MCMC لتجاوز الصعوبات التي تواجهها أساليب التقدير التقليدية. أظهرت هذه الاطروحة كيف يمكن لهذه التقنيات أن تتغلب على صعوبات التقدير في النماذج المعقدة التي تتضمن عدة مستويات من المتغيرات العشوائية. وقد أظهرت النتائج ان استعمال أساليب MCMC يعد أداة فعّالة ومرنة في تقدير النماذج الهرمية متعددة المستويات، وقادرة على التعامل مع تعقيدات وتنوع البيانات بشكل متقدم. [12]

وفي عام 2005، أكد الباحث (Rouder) على الأهمية النظرية والعملية للنماذج الهرمية البيزية في نمذجة التباين متعدد المصادر، إذ بين أن النماذج التقليدية غير الخطية تؤدي إلى تقديرات متحيزة للمعلمات. في المقابل، بين البحث أن النماذج الهرمية البيزية توفر تقديرات متسقة ودقيقة لتأثيرات تلك العناصر. [29]

وفي عام (2006) قدم الباحثان (Browne & Draper) مقارنة منهجية شاملة بين الطرق البيزية والتقليدية المتمثلة بطريقة الإمكان الاعظم، حيث أظهرت النتائج أن الطريقة البيزية هي الأفضل والأكثر قوة ومرونة مع العينات ذات الأحجام الصغيرة أو البيانات المعقدة. بينما تميزت طريقة الإمكان الأعظم بالكفاءة الحسابية لكنها كانت أقل دقة في بعض الحالات المعقدة. هذه المقارنة شكلت إضافة مهمة في فهم مزايا وعيوب كل منهجية. [11]

بينما في عام (2009) استعمل الباحث (Qian) واخرون أسلوب النمذجة الهرمية البيزية لتحليل الظواهر البيئية المعقدة، مثل ظاهرة نقص الأوكسجين في المياه القاعية التي تُعد من القضايا المهمة في علم البيئة البحرية. وقد تمت مقارنة أداء هذه النماذج مع نماذج الانحدار الخطي التقليدية المعتمدة على تحليل التباين (ANOVA). أظهرت النتائج تفوق النمذجة الهرمية البيزية من حيث الدقة الإحصائية وقوة التقدير مقارنةً بتحليل التباين القياسي، ما دفع الباحثين إلى التوصية باعتماد النمذجة الهرمية البيزية كبديل أكثر كفاءة وواقعية عند دراسة الظواهر البيئية البحرية المعقدة . [28]

وفي عام 2014، قدّم الباحث (Chen) واخرون انموذجاً هرمياً بييزياً للتنبؤ بالأمطار والفيضانات باستعمال المتغيرات المناخية، إذ أظهرت نتائج الدراسة قدرة الانموذج على التنبؤ بشكل موثوق ومؤكّد بهطول الامطار الصيفية وتدفقات المياه للموسم المقبل ما يساعد على التقليل والحد من مخاطر الفيضانات. [14]

ففي عام 2017، قدم الباحثان (He & Kolovos) إطار (Bayesian Maximum Entropy) BME المتقدم الذي جمع بشكل مبتكر بين النظرية البيزية ومفاهيم الانتروبي. هذا الإطار (BME) يُعدُّ أداة مرنة لتحليل البيانات في ظل ظروف عدم اليقين واعطاء نتائج ذات دقة عالية على الرغم من نقص المعلومات أحياناً. كما أشارا إلى إمكانية تطبيق BME بشكل واسع على الرغم من الصعوبات الحسابية. [21]

في عام (2019) قام الباحثان (Alaba & Godwin) باستعمال النماذج الهرمية البيزية لتحليل بيانات وفيات الرضع في نيجيريا بهدف الكشف عن عدم تساوي التباينات في معدل الوفيات بين مختلف المناطق الجغرافية داخل البلاد. وقد تم التركيز على تأثير المتغيرات في مستوى الفرد في تفسير خطر وفاة الرضيع، عن طريق بيانات تم جمعها لهذا الغرض وقد توصل الباحث إلى مجموعة من النتائج التي تشير إلى ان 10% يؤثر عليه اختلاف المناطق ما يؤكد أهمية أخذ البُعد المكاني في الحسبان عند تحليل مثل هذه الظواهر. [7]

في عام (2021) قدمت الخزعلي مقارنة بين الطرائق التقليدية مثل طريقة الإمكان الأعظم MLE وطريقة المربعات الصغرى الموزونة $WOLS$ والطرائق الذكية وطريقة الخوارزمية الجينية في نمذجة الانحدار اللوجستي الهرمي، فقد أظهرت التفوق الواضح للشبكات العصبية الاصطناعية ANN . [1]

وفي العام نفسه ركز الباحث (Sosa) وآخرون على تطبيق النمذجة الهرمية البيزية لتحليل البيانات ذات البنية متعددة المستويات وقد أظهرت نتائج الدراسة ان النماذج الهرمية البيزية توفر دقة اعلى ومرونة أكبر مقارنة بالأساليب التقليدية ما يجعلها خياراً ملائماً لتحليل البيانات المعقدة وفهم العلاقات بين المستويات المختلفة. [31]

في عام (2022) استعمل الباحث (Zhang) النمذجة الهرمية البيزية في اطروحة لتحليل البيانات المكانية-الزمانية فائقة التعقيد. وقد أشار إلى ان النمذجة التقليدية تكون غير مجدية وذلك لصعوبة التعامل مع البيانات المعقدة كبيانات انتشار الأمراض أو البيانات الوظيفية المستمرة. إذ استعملت سلاسل ماركوف مونت كارلو ($MCMC$) لتقدير المعلمات في النماذج المعقدة. وقد أظهرت نتائج الاطروحة ان النماذج البيزية الهرمية توفر دقة أكبر في تقدير المعلمات وتسهم بشكل فعال في فهم التبعيات المكانية والزمانية في البيانات. [36]

في عام 2023 أكد الباحث (Veenman et al.) على أهمية إعادة تقييم نمذجة الانحدار الهرمي باستعمال الانموذج البيزي. عالج هذا البحث كيفية التعامل مع البيانات ذات المستويات المتعددة ومعالجة عدم اليقين الملازم لهذا النوع من البيانات ومن الطرائق الحسابية التي تم استعمالها لتحليل النتائج هي سلسلة ماركوف مونت كارلو $MCMC$.. وقد أظهرت النتائج إلى ان النماذج الهرمية البيزية توفر إطار قوي ومرن لتحليل البيانات متعددة المستويات. [34]

اما في عام (2024) قدم الباحث (Dardeno et al.) انموذجاً بييزياً هرمياً لتحليل دوال الاستجابة الترددية (FRF) الخاصة بالأنظمة الديناميكية. استعمل الباحثون أسلوب النمذجة الهرمية البيزية التي تدمج المعلومات عبر مستويات متعددة من البيانات، مع استعمال تقنيات سلسلة ماركوف مونت كارلو ($MCMC$) لتقدير المعلمات. اظهرت نتائج البحث أن هذا الانموذج يوفر تقديرات أكثر دقة واستقراراً مقارنة بالطرائق التقليدية، لاسيما في الحالات التي تتضمن بيانات ضوضاء عالية أو تعقيدات بنوية، ما يعزز القدرة على فهم وتحليل استجابة الأنظمة بشكل موثوق. [16]

وفي عام (2025) قدمت (Halliday) نموذج بيزي هرمي لتصحيح تأخيرات الإبلاغ في رصد الأمراض باستعمال نموذج *Dirichlet*-المعمم-المتعدد. (GDM) إذ أظهرت النتائج تفوق الانموذج من حيث الدقة والكفاءة مقارنة بالأساليب التقليدية، ما يعزز دوره في دعم قرارات الصحة العامة. [20]

على الرغم من التطور الملحوظ في توظيف النمذجة الهرمية البيزية في مجالات متعددة، لا تزال هناك فجوات بحثية جوهرية في تطبيقاتها ضمن القطاع المصرفي. إذ ركزت معظم الدراسات السابقة على المجالات الطبية أو البيئية أو الهندسية، في حين لم يُسجل حتى الآن أي بحث يستثمر بصورة متكاملة القدرات الكامنة في الدمج بين الأسلوب البيزي مع طريقة *MAP*، ومقدر الانتروبي العظمى كقيمة أولية (*initial value*). ضمن نماذج الانحدار اللوجستي ومن هذا المنطلق، جاءت هذه الرسالة لسد هذه الفجوة المعرفية المهمة في المجال المصرفي، عن طريق تطوير انموذج إحصائي متكامل يجمع بين المزايا التحليلية للنمذجة الهرمية البيزية والقوة التفسيرية التي توفرها نظرية الانتروبي.

الفصل الثاني

الجانب النظري

2.1 تمهيد (Preface)

يهدف هذا الفصل إلى استعراض الأسس النظرية والمفاهيم الأساسية ذات الصلة بموضوع الرسالة، مع التركيز على نموذج الانحدار اللوجستي الثنائي المتعدد وطرائق تقدير معالمه، وتحليل البيانات الهرمية أو متعددة المستويات (*Hierarchical or Multilevel Data*)، فضلاً عن عرض أسلوب تقدير رئيسيين هما الأسلوب التقليدي المتمثل بطريقة الامكان الاعظم (MLE) باستعمال خوارزمية نيوتن-رافسون والثاني الأسلوب البيزي بطريقة *MAP (BM)*، فضلاً عن عرض نظرية بيز ومفهوم الإنتروبي وكيفية توظيف مقدر الانتروبي كقيمة أولية مع أسلوب بيز بطريقة *MAP (BME)* في تقدير معالم نموذج الانحدار متعدد المستويات او الهرمي. ويُعد هذا العرض النظري الأساس الذي سيتم البناء عليه في الفصول اللاحقة لتطوير النموذج المقترح وتطبيقه على بيانات البحث.

2.2 الأسس النظرية للانحدار التقليدي:

2.2.1 الانحدار الخطي (Linear regression) [27] ويتضمن الآتي:

1- الانحدار الخطي البسيط (simple linear regression) : يُعد من أبسط نماذج الانحدار الخطي، ويُستعمل لتحليل العلاقة بين متغير معتمد (Dependent Variable) ومتغير توضيحي واحد (Explanatory Variable) ويُفترض أن العلاقة بين المتغيرين خطية، أي يمكن تمثيلها بمعادلة خط مستقيم وعلى النحو الآتي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \quad , \quad i = 1, 2, \dots, n \quad \dots (2 - 1)$$

2- الانحدار الخطي المتعدد (*Multiple linear Regression*) : هو امتداد للانحدار الخطي البسيط، ويُستعمل لنمذجة العلاقة بين متغير معتمد (Dependent Variable) وعدد من المتغيرات التوضيحية (*Explanatory Variables*) يتم التعبير عن النموذج رياضياً على النحو الآتي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad \dots (2 - 2)$$

$$, \quad i = 1, 2, \dots, n$$

2.2.2. الانحدار غير الخطي Non-linear regression: [10]

يعتبر هذا النوع من النماذج غير الخطية والتي تستعمل عندما لا يمكن تمثيل العلاقة بين المتغيرات بخط مستقيم، وإنما تتخذ شكلاً غير خطي، مثل النماذج الأسية، أو اللوغاريتمية، أو اللوجستية. ويُعد الانحدار اللوجستي من النماذج غير الخطية ذات الأهمية الخاصة، وهو محور الاهتمام في هذه الرسالة، ولا سيما صيغة الانحدار اللوجستي الثنائي المتعدد (Multiple Binary Logistic Regression).

2.2.3 الانحدار اللوجستي الثنائي المتعدد [1,5]

(Multiple Binary Logistic Regression)

يُعد نموذج الانحدار اللوجستي (*Logistic Regression Model*) أحد النماذج الإحصائية غير الخطية (*Nonlinear Models*)، ويستعمل لغرض التنبؤ باحتمالية وقوع حدث ما وذلك عن طريق ملاءمة البيانات والمعلومات على المنحني اللوجستي ويعد أحد الأساليب الإحصائية المهمة التي تقوم بتحليل العلاقة بين المتغير التابع النوعي (متغير الاستجابة) الذي يتبع توزيع برنولي (*Bernoulli*) والمتغيرات المستقلة أو التوضيحية $\{x_1, x_2, \dots, x_i\}$ سواء أكانت نوعية أم كمية إذ يكون التعامل مع نماذج الانحدار غير الخطية بحالتين إما بتحويلها إلى الصيغة الخطية لتسهيل عملية التقدير أو البقاء على صيغتها غير الخطية كما هي عند تقدير معالمها. ويُستعمل الانحدار اللوجستي الثنائي عندما يكون للمتغير التابع قيمتان فقط، هما (0,1)، مثل: النجاح مقابل الفشل، أو الإصابة مقابل عدم الإصابة. ويُعبّر عن احتمال حدوث الحالة عندما $(Y = 1)$ بالرمز (φ_i) ، بينما يُمثّل احتمال عدم حدوث الحالة عندما $(Y = 0)$ بالرمز $(1 - \varphi_i)$. بذلك تكون صيغة دالة الكثافة الاحتمالية لتوزيع برنولي كالآتي:

$$P_r(Y_i|X_i) = (\varphi_i)^{y_i}(1 - \varphi_i)^{1-y_i} \quad \dots (2 - 3)$$

اذ ان:

Y_i : متغير تابع ثنائي الاستجابة

φ_i : يمثل احتمال النجاح عند $(Y_i = 1)$

$1 - \varphi_i$: يمثل احتمال الفشل عند ($Y_i = 0$)

لذلك فان توقع المتغير Y_i يمثل احتمال حدوث الاستجابة (النجاح) φ_i أي ان:

$$E(Y_i) = P_r(Y = 1) = \varphi_i \quad \dots (2 - 4)$$

اما تباين المتغير Y_i يكون:

$$Var(Y_i) = \varphi_i(1 - \varphi_i) \quad \dots (2 - 5)$$

أذن صيغة نموذج الانحدار اللوجستي الثنائي المتعدد كالآتي:

$$Y_i = \varphi_i + e_i \quad \dots (2 - 6)$$

إذ ان φ_i تمثل احتمالية النجاح ويمكن ايجادها على وفق الصيغة الآتية:

$$\varphi_i = \frac{e^{(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)}}{(1 + e^{(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)})}$$

$$\varphi_i = \frac{1}{(1 + e^{-(\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + U_i)})} \quad \dots (2 - 7)$$

وعليه الية تحويل الانحدار اللوجستي غير الخطي إلى الصيغة الخطية يكون بطريقتين:

الطريقة الأولى: حاصل ضرب الطرفين في الوسطين للمعادلة (2 - 7) ثم اخذ اللوغاريتم الطبيعي للطرفين للوصول إلى الصيغة الخطية وكالآتي: [15]

$$1 = \varphi_i + \varphi_i e^{-(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)}$$

$$1 - \varphi_i = \varphi_i e^{-(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)}$$

$$\text{Log}(1 - \varphi_i) = \text{Log}(\varphi_i e^{-(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)})$$

$$\text{Log}(1 - \varphi_i) = \text{Log}\varphi_i - (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i)$$

$$\text{Log}\varphi_i - \text{Log}(1 - \varphi_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i$$

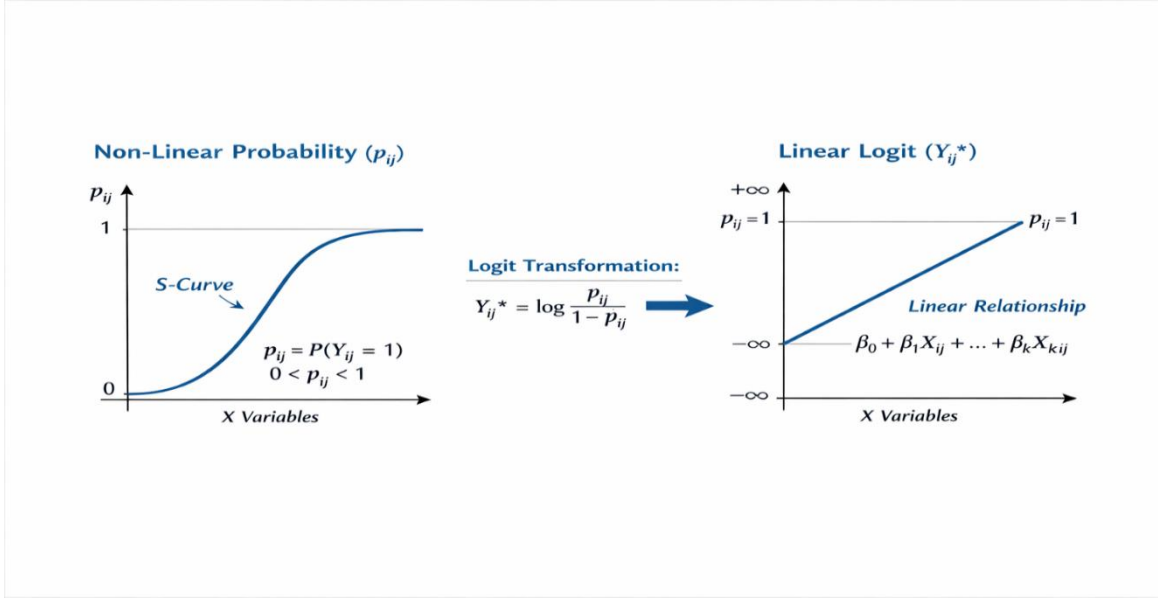
$$\text{Log} \left[\frac{\varphi_i}{1 - \varphi_i} \right] = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i \quad \dots (2 - 8)$$

$$Y_i^* = \text{Log} \left[\frac{\varphi_i}{1-\varphi_i} \right] \quad \text{لنفترض ان}$$

عليه يتم إعادة كتابة الصيغة أعلاه بالشكل الآتي:

$$Y_i^* = \beta_0 + \sum_{k=1}^K \beta_k X_{ki} + U_i \quad \dots (2 - 9)$$

$$k = 1, \dots, k \quad , i = 1, 2, \dots, n$$



الشكل (1-2) يوضح تحويل الانموذج من الصيغة الغير خطية الى الصيغة الخطية

دالة اللوجيت (Logit function) هي دالة رياضية تستعمل لتحويل احتمال يتراوح بين (0 و 1) إلى قيمة تمتد من $-\infty$ إلى ∞ حيث تعبر φ_i عن احتمال ضمن الفترة $0 \leq \varphi_i \leq 1$ والصيغة (2 - 9) تمثل الصيغة الخطية لانموذج الانحدار اللوجستي.

الطريقة الثانية: أخذ اللوغاريتم الطبيعي لنسبة احتمال النجاح إلى احتمال الفشل، وكالاتي: [1]

$$Y_i^* = \text{Logit}(\varphi_i) = \text{Log} \left[\frac{\varphi_i}{1 - \varphi_i} \right]$$

$$Y_i^* = \text{Logit}(\varphi_i) = \text{Log} \left[\frac{e^{(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + U_i)}}{1 + e^{(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + U_i)}} \right]$$

$$Y_i^* = \text{Logit}(\varphi_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i$$

$$Y_i^* = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + U_i$$

$$Y_i^* = \beta_0 + \sum_{k=1}^K \beta_k X_{ki} + U_i$$

وعليه سيتم عرض نموذج الانحدار اللوجستي للبيانات الهرمية بالصيغة الخطية.

2.3 النماذج متعددة المستويات والبيانات الهرمية:

(Multilevel Models and Hierarchical Data) [4,19]

هي نوع من النماذج الإحصائية المصممة لتحليل البيانات المنظمة في طبقات أو مستويات، مثل الأفراد داخل صفوف دراسية، أو الموظفين ضمن مؤسسات. وتتميز هذه النماذج بقدرتها على التمييز بين التباين على مستوى الفرد والتباين على مستوى المجموعة، ما يسمح بتقديم تفسير أكثر دقة وشمولاً للعلاقات الإحصائية بين المتغيرات. في العديد من الدراسات، يتم استعمال نموذج بمستويين فقط لتجنب التعقيد الناتج عن تحليل مستويات متعددة. في هذا النوع من النماذج:

المستوى الأول يمثل الوحدات الفردية، مثل الطلاب أو المرضى، والتي تُقاس عندها المتغيرات المستقلة الخاصة بالفرد، مثل العمر، الجنس، أو عدد ساعات الدراسة. هذه المتغيرات تؤثر بشكل مباشر في المتغير التابع، مثل حالة النجاح أو الإصابة بمرض معين.

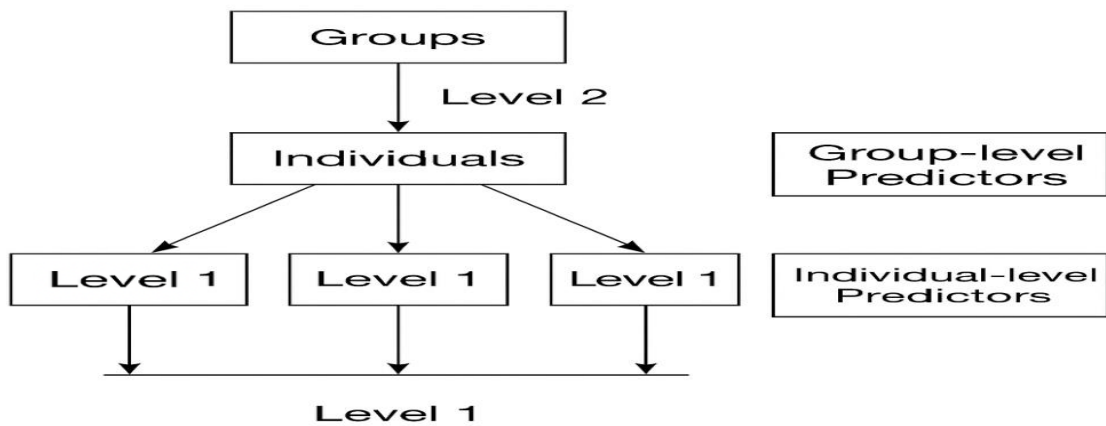
أما المستوى الثاني فيشير إلى المجموعات التي ينتمي إليها الأفراد (مثل الصفوف أو المستشفيات)، ويشمل متغيرات تُقاس على مستوى المجموعة كنوع المؤسسة أو موقعها أو حجمها. هذه المتغيرات يمكن أن تؤثر في المتغير التابع بصورة مباشرة، أو بشكل غير مباشر عن طريق تعديل أثر متغيرات المستوى الأول. ومن ثمّ، يسمح هذا النوع من النماذج بفصل التأثيرات الفردية عن تأثيرات المجموعات، مع إمكانية إدخال معلمات عشوائية تعكس الفروق بين المجموعات، ما يزيد من دقة التفسير ومرونة الانموذج. وقد تنشأ أيضاً تفاعلات بين متغيرات المستويين، وهو ما يُبرز أهمية استعمال انموذج انحدار لوجستي متعدد المستويات. في الاسلوب التقليدي، تُقدّر المعلمات بشكل منفصل لكل مستوى، بينما في الاسلوب البيزي تُرتب المعلمات ضمن هيكل هرمي من التوزيعات الأولية (Hierarchical Priors Distributions)،

الفصل الثاني الجانب النظري

بحيث يعكس المستوى الأول التباين داخل المجموعات، ويعكس المستوى الثاني التباين بين المجموعات.

وعند تحليل البيانات الهرمية أو متعددة المستويات باستعمال نموذج الانحدار اللوجستي، يُعد إدخال التأثيرات العشوائية (Random Effects) أمرًا أساسيًا لتمثيل التباين المشترك بين المجموعات. تساعد هذه التأثيرات على توصيف الفروق غير المفسرة بين الوحدات على المستوى الأعلى (المستوى الثاني أو أكثر)، ما يعزز دقة النماذج ومرونتها.

وعلى الرغم من أن مصطلحي البيانات متعددة المستويات (Multilevel Data) والبيانات الهرمية (Hierarchical Data) يشيران في جوهرهما إلى نفس البنية التراتبية، فإن الأدبيات الإحصائية تميل إلى استعمال "البيانات متعددة المستويات" في المفهوم التقليدي (Frequentist)، بينما يُستعمل مصطلح "البيانات الهرمية" في المفهوم البيزي (Bayesian) لوصف النماذج التي تعتمد على توزيعات أولية هرمية (Hierarchical Priors) وفي سياق هذه الرسالة، سيتم اعتماد مصطلح البيانات الهرمية عند عرض نموذج الانحدار اللوجستي، سواء في صياغته التقليدية أم البيزية، نظرًا لارتباطه المباشر بمحور الرسالة. والشكل الآتي يوضح ما تقدم.



شكل (2-2) شكل يوضح الانموذج المتعدد المستويات بمستويين

2.4 نموذج الانحدار اللوجستي (بمستويين)

Two-level Logistic Regression Model [32,35,8]

يُعدُّ من أحد النماذج الإحصائية التي تستعمل عندما تكون البيانات منظمة بشكل هرمي، إذ يكون لدينا وحدات فردية أصغر (مثل الافراد) ضمن مجموعات أكبر (مثل الصفوف او المستشفيات) ويهدف هذا الانموذج إلى تقدير احتمال وقوع حدث ثنائي، مع الأخذ بعين الاعتبار التباين داخل المجموعات والتباين بين المجموعات. يمثل المستوى الأول (Level 1) المتغيرات المستقلة الخاصة بالأفراد (مثل العمر، الجنس)، وهي متغيرات تختلف من فرد لآخر داخل كل مجموعة، وتؤثر مباشرة في المتغير التابع. أما المستوى الثاني (Level 2)، فيتضمن المتغيرات المستقلة التي تمثل خصائص المجموعات التي ينتمي إليها الأفراد (مثل خصائص المدارس، أو المناطق الجغرافية)، وهذه المتغيرات تؤثر إما بشكل مباشر في المتغير التابع، عن طريق تأثيرها العام في جميع الأفراد في المجموعة أو بشكل غير مباشر عن طريق تعديل العلاقة بين متغيرات المستوى الأول والمتغير التابع، وهو ما يعرف بـ التآثيرات التفاعلية (*cross-level interactions*). وبما أن الرسالة تعتمد على انموذج انحدار لوجسـتي بمستويين (Two-level Logistic Regression Model)، سيتم عرض الصيغة العامة لأنموذج الانحدار متعدد المستويات (بمستويين) وفق الاسلوب التقليدي. وكالاتي:

$$Y_{ij}^* = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \dots + \beta_{kj}X_{kij} + e_{ij} \quad \dots (2 - 10)$$

الانموذج (2-10) يمثل (Level 1) ويمكن كتابته بشكل اخر وكالاتي:

$$Y_{ij}^* = \beta_{0j} + \sum_{k=1}^K \beta_{kj}X_{kij} + e_{ij} \quad \dots (2 - 11)$$

$k = 1, 2, \dots, K$: عدد المتغيرات التوضيحية عند المستوى الاول

$i = 1, 2, \dots, n_j$: عدد الافراد داخل المجموعة j

$j = 1, 2, \dots, L$: عدد المجموعات الكلية في المستوى الثاني

Y_{ij}^* : يمثل المتغير المعتمد للملاحظة i عند المستوى الأول للمجموعة j

X_{kij} : تمثل المتغير التوضيحي k للفرد i عند المستوى الأول للمجموعة j

معلمات انموذج الانحدار اللوجستي الخطي عند المستوى الأول

$\beta_{0j}, \beta_{1j}, \dots, \beta_{kj}$: يمثل الخطأ العشوائي للمستوى الأول للفرد i في المجموعة j

الان يتم كتابة الانموذج الذي يمثل المستوى الثاني (**Level 2**) وكالاتي:

$$\left. \begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + \dots + \gamma_{0M}W_{Mj} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}W_{1j} + \gamma_{12}W_{2j} + \dots + \gamma_{1M}W_{Mj} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}W_{1j} + \gamma_{22}W_{2j} + \dots + \gamma_{2M}W_{Mj} + u_{2j} \\ &\vdots \\ \beta_{kj} &= \gamma_{k0} + \gamma_{k1}W_{1j} + \gamma_{k2}W_{2j} + \dots + \gamma_{kM}W_{Mj} + u_{kj} \end{aligned} \right\} \dots (2 - 12)$$

الصيغة (2 - 12) تمثل (**Level 2**) ويمكن كتابتها بشكل اخر وكالاتي

$$\beta_{0j} = \gamma_{00} + \sum_{m=1}^M \gamma_{0m}W_{mj} + u_{0j} \quad \dots (2 - 13)$$

$$\beta_{kj} = \gamma_{k0} + \sum_{m=1}^M \gamma_{km}W_{mj} + u_{kj} \quad \dots (2 - 13)$$

$m = 1, 2, \dots, M$: عدد المتغيرات التوضيحية عند المستوى الثاني

W_{mj} : تمثل المتغير التوضيحي m عند المستوى الثاني للمجموعة j

معلمات انموذج الانحدار الخطي عند المستوى الثاني : $\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11}, \dots, \gamma_{kM}$

γ_{00}, γ_{k0} : التأثيرات الثابتة (**fixed Effects**) للحد الثابت والميول في الانموذج.

γ_{0m}, γ_{km} : معلمات تأثير المتغيرات المستقلة في المستوى الثاني ؛ أي مقدار تأثير كل متغير

توضيحي W_{mj} على الحد الثابت او الميل.

u_{0j} : يمثل التأثير العشوائي (**Random Effect**) للحد الثابت للمجموعة j ؛

u_{kj} : تمثل التأثيرات العشوائية (**Random Effects**) لمعاملات الميول (**Slopes**) عبر

المجموعات.

الفصل الثاني ————— الجانب النظري

عند تعويض معادلات المستوى الثاني في معادلة المستوى الأول، نحصل على المعادلة الكلية التي تضم جميع المتغيرات والتأثيرات عند المستويين، وتُستعمل في تقدير معاملات الانموذج. عن طريق تعويض المعادلات الموضحة في (2 - 12) في المعادلة (2 - 10) وكالاتي:

$$Y_{ij}^* = (\gamma_{00} + \sum_{m=1}^M \gamma_{0m} W_{mj} + u_{0j}) + (\gamma_{10} + \sum_{m=1}^M \gamma_{1m} W_{mj} + u_{1j}) X_{1ij} + (\gamma_{20} + \sum_{m=1}^M \gamma_{2m} W_{mj} + u_{2j}) X_{2ij} + \dots + (\gamma_{k0} + \sum_{m=1}^M \gamma_{km} W_{mj} + u_{kj}) X_{kij} + e_{ij} \quad \dots (2 - 14)$$

بالتبسيط نحصل على:

$$Y_{ij}^* = \gamma_{00} + \sum_{m=1}^M \gamma_{0m} W_{mj} + \gamma_{10} X_{1ij} + \sum_{m=1}^M \gamma_{1m} W_{mj} X_{1ij} + \gamma_{20} X_{2ij} + \sum_{m=1}^M \gamma_{2m} W_{mj} X_{2ij} + \dots + \gamma_{k0} X_{kij} + \sum_{m=1}^M \gamma_{km} W_{mj} X_{kij} + u_{0j} + u_{1j} X_{1ij} + u_{2j} X_{2ij} + \dots + u_{kj} X_{kij} + e_{ij}$$

$$Y_{ij}^* = \gamma_{00} + \sum_{k=1}^k \gamma_{k0} X_{kij} + \sum_{m=1}^M \gamma_{0m} W_{mj} + \sum_{m=1}^M \sum_{k=1}^k \gamma_{km} W_{mj} X_{kij} + \sum_{k=1}^k u_{kj} X_{kij} + u_{0j} + e_{ij} \quad \dots (2 - 15)$$

$\sum_{k=1}^k \gamma_{k0} X_{kij}$: تمثل تأثير متغيرات المستوى الأول (Level 1) على المتغير المعتمد .

$\sum_{m=1}^M \gamma_{0m} W_{mj}$: تمثل تأثير متغيرات المجموعة (Level 2) على الحد الثابت (intercept)

$\sum_{m=1}^M \sum_{k=1}^k \gamma_{km} W_{mj} X_{kij}$: تمثل التفاعل بين متغيرات المستوى الأول والثاني (Cross Level Interactions).

$\sum_{k=1}^k u_{kj} X_{kij}$: تمثل التأثير العشوائي للميول (slopes) بين المجموعات.

$$\underbrace{\gamma_{00} + \sum_{k=1}^k \gamma_{k0} X_{kij} + \sum_{m=1}^M \gamma_{0m} W_{mj} + \sum_{m=1}^M \sum_{k=1}^k \gamma_{km} W_{mj} X_{kij}}_{\text{fixed part}}$$

تمثل الجزء الثابت من الانموذج؛ تبدأ بالحد الثابت العام؛ ومن ثم تأثير المتغيرات الفردية؛ ومن ثم تأثير خصائص المجموعة؛ واخيراً تأثير التفاعل بين الفرد والمجموعة.

$$\text{يمثل هذا الجزء مكونات التباين العشوائي، سواء الناتج عن } \underbrace{\sum_{k=1}^k u_{kj} X_{kij} + u_{0j} + e_{ij}}_{\text{random part}}$$

اختلاف المجموعات عن بعضها البعض أو الناتج عن التباين بين الأفراد داخل نفس المجموعة.

2.5 طرق التقدير التقليدية: Traditional Estimation Method

2.5.1 طريقة الإمكان الأعظم

Maximum Likelihood Method (MLE)

تعدُّ طريقة *MLE* من أكثر الطرق استعمالاً لتقدير المعلمات في نماذج الانحدار الخطي وغير الخطي وذلك عن طريق تعظيم دالة الاحتمال للمتغير المعتمد (الاستجابة) مع استعمال خوارزمية نيوتن-رافسون لتكرار التقدير حتى الوصول إلى القيمة المثلى. وغايتها الحصول على المعادلة التقديرية لأنموذج الانحدار التي تكون على أكثر من مستوى واحد أي بمستويات متعددة للوصول إلى أفضل التنبؤات عن الظاهرة المدروسة إذ تفترض هذه الطريقة أن لمتغير الاستجابة توزيعاً معلوماً، وعن طريق هذا الافتراض يمكن بناء دالة الإمكان الأعظم لعينة مكونة من n من المشاهدات.. لذا يمكن تطبيقها مع الانحدار اللوجستي لنمذجة العلاقة بين متغير الاستجابة الثنائي والمتغيرات التوضيحية ما يسمح بتقدير المعلمات بشكل أفضل وكالاتي. [1,17]

$$Y_{ij} \sim \text{Bernoulli}(\varphi(x_{ij}))$$

إذ ان Y_{ij} هو متغير استجابة يأخذ قيما (0 و1) لذا يمكن كتابة دالة الإمكان وكالاتي:

$$L(\beta) = \prod_{i=1}^n \prod_{j=1}^J (\varphi(x_{ij}))^{Y_{ij}} (1 - \varphi(x_{ij}))^{1-Y_{ij}} \quad \dots (2 - 16)$$

$$i = 1, 2, \dots, n, \quad j = 1, 2, \dots, J$$

$$\frac{\varphi(x_{ij})}{1 - \varphi(x_{ij})} = e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}$$

$$1 - \varphi(x_{ij}) = \frac{1}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}$$

$$\log \frac{\varphi(x_{ij})}{1 - \varphi(x_{ij})} = (\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij}) \text{ or } X_{ij}^T \beta$$

إذ ان

$$\varphi(x_{ij}) = \frac{e^{X_{ij}^T \beta}}{1 + e^{X_{ij}^T \beta}}, \quad X_{ij}^T \beta = [X_{i0} \quad X_{i1} \quad \dots \quad X_{ik}] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$$

يمكن اعادة كتابة دالة الإمكان الأعظم للصيغة (2-16) وكالاتي:

$$L(\beta) = \prod_{i=1}^n \prod_{j=1}^J (e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^{Y_{ij}} \left(\frac{1}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} \right) \dots \quad (2-17)$$

وبأخذ اللوغاريتم لدالة الإمكان (2-17) يتم الحصول على الاتي:

$$\log L(\beta)$$

$$= \log \left[\prod_{i=1}^n \prod_{j=1}^J (e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^{Y_{ij}} \frac{1}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} \right] \dots \quad (2-18)$$

$$\log L(\beta) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} \log(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) + \log(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^{-1}] \dots \quad (2-19)$$

$$\log L(\beta) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} (\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij}) - \log(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})] \dots \quad (2-20)$$

وللحصول على تقديرات معاملات نموذج الانحدار نقوم بعملية الاشتقاق للمعلمة β_0 وللصيغة (2-20) وكالاتي:

$$\frac{\partial \log L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \sum_{j=1}^J \left[Y_{ij} - \frac{e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} * 1 \right] \quad \dots (2-21)$$

$$\frac{\partial \log L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] = 0 \quad \dots (2-22)$$

اما المشتقة الأولى ل: β_1 للصيغة (2-20) تكون كالآتي:

$$\frac{\partial \log L(\beta)}{\partial \beta_1} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} X_{1ij} - \frac{e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} * X_{1ij}] \quad \dots (2-23)$$

$$\frac{\partial \log L(\beta)}{\partial \beta_1} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} X_{1ij} - \varphi(x_{ij}) X_{1ij}] = 0$$

$$\frac{\partial \log L(\beta)}{\partial \beta_1} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij}) X_{1ij}] = 0 \quad \dots (2-24)$$

⋮

في حين المشتقة الأولى ل: β_k للصيغة (2-20) تكون كالآتي:

$$\frac{\partial \log L(\beta)}{\partial \beta_k} = \sum_{i=1}^n \sum_{j=1}^J \left[Y_{ij} X_{kij} - \frac{e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} * X_{kij} \right] \quad \dots (2-25)$$

$$\frac{\partial \log L(\beta)}{\partial \beta_k} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} X_{kij} - \varphi(x_{ij}) X_{kij}] = 0$$

$$\frac{\partial \log L(\beta)}{\partial \beta_k} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij}) X_{kij}] = 0 \quad \dots (2-26)$$

نلاحظ من المعادلة (2-22) و (2-24) و (2-26) انه من الصعب الوصول إلى صيغ تقدير محددة تحليلياً لذا يتم استعمال طرائق تقدير تكرارية كطريقة نيوتن رافسون التكرارية. ومن خطوات تطبيق هذه الطريقة هي تقديم قيم تقدير أولية (*initial estimated value*) عن المعلمات التي يتم افتراضها من لدن الباحث وتكتب بشكل $\beta_{(0)}$.

$$\beta_{(0)} = [\beta_{(0)0}, \beta_{(0)1}, \dots, \beta_{(0)k}]'$$

ويتم استعمال القيمة الأولية كإدخال لاستخراج المقدر للتكرار الثاني $\beta_{(r+1)}$ وهذا بدوره يستخرج تكرار المعلمات بالتكرار الثاني. فإذا كان الفرق بين التكرار الثاني والقيمة الأولية التي تم إدخالها مساوياً تقريباً إلى قيمة ϵ التي تم تقديرها من لدن الباحث وهي قيمة موجبة وصغيرة جداً هنا اتوقف واعد β_1 هو التقدير اما إذا كان حاصل الفرق أكبر من قيمة ϵ استمر واعد التكرار الثاني هو القيمة الأولية وهكذا اما صيغة التقدير التكرارية لنيوتن رافسون هي كالآتي:

$$\beta_{ML(r+1)} = \beta_r - H(\beta_r)^{-1}U(\beta_r) \quad \dots (2 - 27)$$

$$U(\beta) = \begin{bmatrix} \frac{\partial l(\beta)}{\partial \beta_0} \\ \frac{\partial l(\beta)}{\partial \beta_1} \\ \vdots \\ \frac{\partial l(\beta)}{\partial \beta_k} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] \\ \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})X_{1ij}] \\ \vdots \\ \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})X_{kij}] \end{bmatrix}$$

تعتمد طريقة نيوتن رافسون على حساب مصفوفة هيس وكالآتي:

$$H(\beta)^{-1} = \begin{bmatrix} \frac{\partial^2 l(\beta)}{\partial \beta_0^2} & \frac{\partial^2 l(\beta)}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 l(\beta)}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 l(\beta)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l(\beta)}{\partial \beta_1^2} & \dots & \frac{\partial^2 l(\beta)}{\partial \beta_1 \partial \beta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_0} & \frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_1} & \dots & \frac{\partial^2 l(\beta)}{\partial \beta_k^2} \end{bmatrix}^{-1}$$

اذن المشتقة الثانية بالنسبة الى β_0 في مصفوفة هيس للمعادلة (2-22) تكون كالآتي:

$$\frac{\partial^2 \log L(\beta)}{\partial (\beta_0^2)} = 0 - \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) (e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\ \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) (0 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) 1}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right]$$

$$\begin{aligned}
 &= - \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= - \left[\frac{e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} - \left(\frac{e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}}{1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}} \right)^2 \right] \\
 &= - [\varphi(x_{ij}) - (\varphi(x_{ij}))^2] \\
 &= -\varphi(x_{ij})[1 - \varphi(x_{ij})] \quad \dots (2 - 28)
 \end{aligned}$$

where $\varphi(x_{ij}) = \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}$

المشتقة الثانية بالنسبة الى β_1 في مصفوفة هيس للمعادلة (2-24) تكون كالاتي:

$$\begin{aligned}
 \frac{\partial^2 \log L(\beta)}{\partial (\beta_1^2)} &= 0 \\
 &- X_{1ij} \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) X_{1ij}}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(0 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})}) X_{1ij}}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{1ij}^2 \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{1ij}^2 \left[\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} - \left(\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} \right)^2 \right] \\
 &= -X_{1ij}^2 [\varphi(x_{ij}) - (\varphi(x_{ij}))^2] \\
 &= -X_{1ij}^2 \varphi(x_{ij})[1 - \varphi(x_{ij})] \quad \dots (2 - 29) \\
 &\vdots
 \end{aligned}$$

المشتقة الثانية بالنسبة الى β_K في مصفوفة هيس للمعادلة (2-26) تكون كالاتي:

$$\begin{aligned}
 \frac{\partial^2 \log L(\beta)}{\partial (\beta_k^2)} &= 0 \\
 &- X_{kij} \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(0 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})X_{kij}}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{kij}^2 \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{kij}^2 \left[\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} - \left(\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} \right)^2 \right] \\
 &= -X_{kij}^2 [\varphi(x_{ij}) - (\varphi(x_{ij}))^2] \\
 &= -X_{kij}^2 \varphi(x_{ij}) [1 - \varphi(x_{ij})] \quad \dots (2 - 30)
 \end{aligned}$$

اما المشتقة الثانية الجزئية بالنسبة الى β_1, β_0 في مصفوفة هيس تكون كالآتي:

$$\begin{aligned}
 \frac{\partial^2 \log L(\beta)}{\partial \beta_0, \beta_1} &= 0 - \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})X_{1ij}}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(0 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})X_{1ij}}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{1ij} \left[\frac{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right. \\
 &\quad \left. - \frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})^2} \right] \\
 &= -X_{1ij} \left[\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} - \left(\frac{(e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})}{(1 + e^{(\beta_0 + \beta_1 x_{1ij} + \dots + \beta_k x_{kij})})} \right)^2 \right] \\
 &= -X_{1ij} [\varphi(x_{ij}) - (\varphi(x_{ij}))^2]
 \end{aligned}$$

$$= -X_{1ij} \varphi(x_{ij})[1 - \varphi(x_{ij})] \quad \dots (2 - 31)$$

وبنفس الأسلوب يمكن الحصول على المشتقات الجزئية الأخرى الموجودة في مصفوفة معكوس هيسين.

$$H(\beta) = - \begin{bmatrix} \sum_{i=1}^n \sum_{j=1}^J [\varphi(x_{ij})[1 - \varphi(x_{ij})]] & \sum_{i=1}^n \sum_{j=1}^J [X_{1ij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] & \dots & \sum_{i=1}^n \sum_{j=1}^J [X_{kij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] \\ \sum_{i=1}^n \sum_{j=1}^J [X_{1ij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] & \sum_{i=1}^n \sum_{j=1}^J [X_{1ij}^2 \varphi(x_{ij})[1 - \varphi(x_{ij})]] & \dots & \sum_{i=1}^n \sum_{j=1}^J [X_{1ij} X_{kij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n \sum_{j=1}^J [X_{kij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] & \sum_{i=1}^n \sum_{j=1}^J [X_{kij} X_{1ij} \varphi(x_{ij})[1 - \varphi(x_{ij})]] & \dots & \sum_{i=1}^n \sum_{j=1}^J [X_{kij}^2 \varphi(x_{ij})[1 - \varphi(x_{ij})]] \end{bmatrix}^{-1}$$

وأخيرا هذه هي الخطوات التفصيلية التي تم ذكرها لتقدير معالم نموذج الانحدار اللوجستي باستعمال طريقة الإمكان الأعظم بالاعتماد على خوارزمية نيوتن- رافسون.

2.5.2 خطوات خوارزمية نيوتن- رافسون لتقدير معالم الانموذج بطريقة الإمكان الأعظم (MLE):

الخطوة الأولى: التهيئة الأولية

1- اختيار قيم ابتدائية للمعاملات $\beta_{(0)} = [\beta_{(0)0}, \beta_{(0)1}, \dots, \beta_{(0)k}]^T$

2- تحديد قيمة (ϵ) والتي تكون عادة 10^{-6}

3- تحديد الحد الأقصى لعدد التكرارات (*MaxIter*) عادة 500-1000

الخطوة الثانية: العمليات التكرارية لتقدير المعالم لكل $r = 0, 1, 2, \dots, \text{MaxIter}$ وكما يأتي:

1. حساب الاحتمالات المتوقعة:

$$\varphi_{(r)}(x_{ij}) = \frac{e^{X_{ij}^T \beta_{(r)}}}{1 + e^{X_{ij}^T \beta_{(r)}}}$$

2. حساب الاشتقاق الأول لدالة الإمكان اللوغاريتمية:

$$U(\beta_{(r)}) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi_{(r)}(x_{ij})] X_{ij}$$

3. حساب مصفوفة هيسين (الاشتقاق الثاني):

$$H(\beta_{(r)}) = - \sum_{i=1}^n \sum_{j=1}^J \varphi_{(r)}(x_{ij}) [1 - \varphi_{(r)}(x_{ij})] X_{ij} X_{ij}^T$$

4. معادلة التقدير التكرارية:

$$\beta_{(r+1)} = \beta_{(r)} - [H(\beta_{(r)})]^{-1}U(\beta_{(r)})$$

الخطوة الثالثة: اختبار شرط التقارب

- 1- إذا كان $\|\beta_{(r+1)} - \beta_{(r)}\| < \varepsilon$ انتقل إلى الخطوة 4
 وإلا إذا كان $r = \text{MaxIter}$ أي إذا بلغ عدد التكرارات الحد الأقصى المحدد مسبقاً من دون تحقيق شرط التقارب، أتوقف وأبلغ عن عدم التقارب
 • وإلا: كرر الخطوة الثانية مع $r = r + 1$

الخطوة الرابعة: النتائج

- المعلمات المقدرة $\beta_{MLE} = \beta_{(r+1)}$
- مصفوفة التباين $Cov(\beta_{MLE}) = -[H(\beta_{MLE})]^{-1}$
- أخطاء معيارية $SE(\beta_k) = \sqrt{\text{diag}(Cov(\beta_{MLE}))_k}$

2.6 الاسلوب البيزي بطريقة MAP : [6,9]

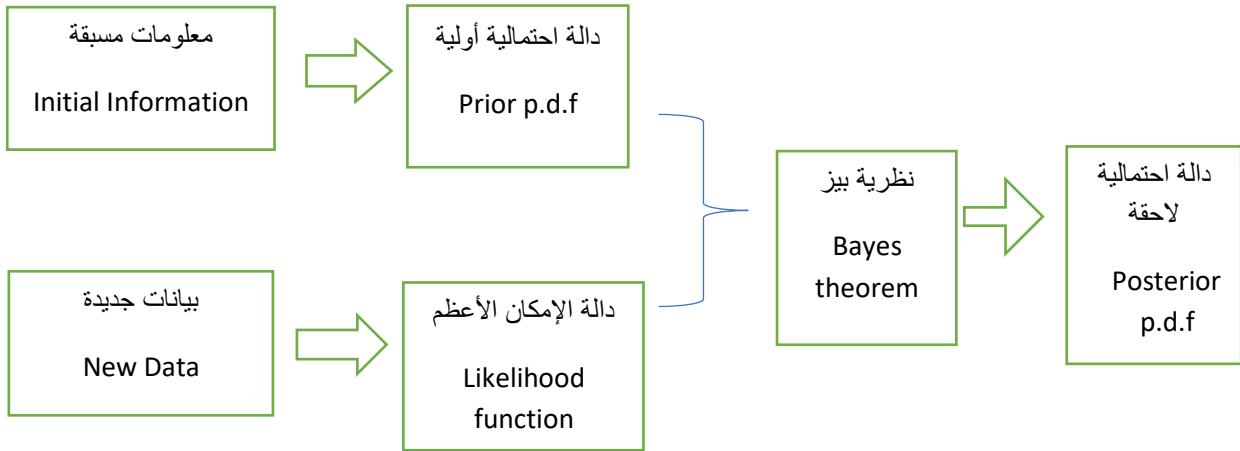
The Bayesian method in a MAP way

يتكون انموذج الانحدار اللوجستي الخطي متعدد المستويات على وفق المفهوم البيزي (*Bayesian*) عندما تحدث التأثيرات العشوائية (*Random Effects*) بامتلاك معلمات المستوى الأول توزيعات أولية (*prior distributions*) وبمعلمات معلومة والتي بالاعتماد عليها وبيانات العينة الحالية يتم تطبيق النظرية البيزية للحصول على الدوال اللاحقة (*Posterior functions*) وبالاعتماد على طريقة (*Maximum A Posterior*) كما أوضح ذلك [13] في سياق النمادج الهرمية البيزية (*Bayesian Hierarchical Models*) التي تسمح بدمج المعلومات المسبقة مع بيانات العينة عبر مستويات متعددة.

قبل عرض تقدير معلمات الانموذج لابد من توضيح مفهوم نظرية بيز: [3]

في النمادج البيزية يُنظر إلى معلمات الانموذج $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ المراد تقديرها على انها متغيرات عشوائية لها توزيعات أولية (*Prior Distribution*) تعكس المعرفة المسبقة. وعند توفر بيانات جديدة، يتم تحديث هذه المعرفة بدمجها مع دالة الإمكان (*Likelihood Function*) للوصول إلى التوزيع اللاحق، (*Posterior Distribution*) الذي يمثل الأساس لتقديرات بيز (*Bayesian Estimators*) لأي معلمة باستعمال دالة الخسارة (*Loss*)

(Function) المختارة، مثل دالة الخسارة التربيعية (Quadratic Loss Function). يمكن تلخيص العملية عن طريق المخطط الآتي:



مخطط (2-3) مراحل دالة الكثافة الاحتمالية اللاحقة [2,3]

ومن الجدير بالذكر ان النمذجة البيزية تعد أسلوباً احصائياً يهدف الى تحديث المعلومات السابقة عن ظاهرة معينة بوجود البيانات الجديدة للحصول على التوزيع اللاحق posterior Distribution لذا فالتوزيع الاولي (prior distribution) لمعلمت انموذج الانحدار اللوجستي يكون لكل معلمة من معلمت الانموذج. أما في النماذج الهرمية، فإن معلمت المستوى الثاني يمكن أن تُعامل كمتغيرات عشوائية أو معتمدة. ويمكن تصنيف التوزيعات الأولية (Prior Distributions) المستعملة لتقدير معلمت انموذج الانحدار اللوجستي إلى أربعة أنواع رئيسة وكالاتي:

1-دالة الكثافة الاحتمالية الأولية غير المعلوماتية Non Informative prior pdf

2-دالة الكثافة الاحتمالية الأولية المعلوماتية Informative prior pdf

3-دالة الكثافة الأولية المرافقة الطبيعية Normal Conjugate Prior pdf

4-دالة الكثافة الأولية المعتمدة على عينات سابقة Based on Previous Samples Prior pdf

وبقدر تعلق الرسالة، سيتم الاقتصار على استعمال دالة الكثافة الاحتمالية الأولية المرافقة ذات التوزيع الطبيعي لمعاملات نموذج الانحدار اللوجستي. وبناءً على ذلك، يتم الحصول على التوزيع اللاحق لمعاملات النموذج بالاعتماد على دالة الإمكان والتوزيع الأولي المفترض، وفق الاسلوب البيزي، كما سيتم توضيحه في الصيغ اللاحقة.

يتم تقدير النموذج الانحدار اللوجستي لمتعدد المستويات بواسطة أسلوب بيز بطريقة MAP كالآتي:

[13,23]

$$L(\beta) = \prod_{i=1}^n \prod_{j=1}^J (\varphi(x_{ij}))^{Y_{ij}} (1 - \varphi(x_{ij}))^{1-Y_{ij}} \quad \dots (2 - 32)$$

$$L(\beta) = \prod_{i=1}^n \prod_{j=1}^J \left(\frac{\varphi(x_{ij})}{1 - \varphi(x_{ij})} \right)^{Y_{ij}} 1 - \varphi(x_{ij}) \quad \dots (2 - 33)$$

$$i = 1, 2, \dots, n, \quad j = 1, 2, \dots, J$$

$$\frac{\varphi(x_{ij})}{1 - \varphi(x_{ij})} = e^{X_{ij}^T \beta}, \quad 1 - \varphi(x_{ij}) = \frac{1}{1 + e^{X_{ij}^T \beta}} \quad \text{إذ أن}$$

وعليه يمكن ان نبسط الصيغة (2-32) لتصبح كالآتي:

$$L(\beta) = \prod_{i=1}^n \prod_{j=1}^J (e^{X_{ij}^T \beta})^{Y_{ij}} \frac{1}{1 + e^{X_{ij}^T \beta}} \quad \dots (2 - 34)$$

عند أخذ اللوغاريتم لدالة الإمكان للصيغة (2-34) يتم الحصول على الآتي:

$$\log L(\beta) = \log \left[\prod_{i=1}^n \prod_{j=1}^J (e^{X_{ij}^T \beta})^{Y_{ij}} \frac{1}{1 + e^{X_{ij}^T \beta}} \right] \quad \dots (2 - 35)$$

$$\log L(\beta) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} \log(e^{X_{ij}^T \beta}) + \log(1 + e^{X_{ij}^T \beta})^{-1}] \quad \dots (2 - 36)$$

$$\log L(\beta) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} X_{ij}^T \beta - \log(1 + e^{X_{ij}^T \beta})] \quad \dots (2 - 37)$$

نفترض ان التوزيع الاولي للمعلمات يتبع توزيعاً طبيعياً قياسياً.

$$\beta \sim N(0, \sigma^2)$$

وبالاعتماد على دالة الكثافة الاحتمالية الأولية المرافقة لمتجه β وبافتراض ثبوت قيمة σ^2 هو 1

نحصل على التوزيع الاولي وكالاتي:

$$\pi(\beta) \propto e^{\left\{-\frac{1}{2}\beta^T \Sigma^{-1} \beta\right\}} \quad \dots (2-38)$$

وعليه فإن التوزيع اللاحق يكتب كالاتي:

$$\pi(\beta|x) \propto L(\beta)\pi(\beta) \quad \dots (2-39)$$

$$\pi(\beta|x) \propto \prod_{i=1}^n \prod_{j=1}^J (e^{X_{ij}^T \beta})^{Y_{ij}} \frac{1}{1 + e^{X_{ij}^T \beta}} * e^{\left\{-\frac{1}{2}\beta^T \Sigma^{-1} \beta\right\}} \quad \dots (2-40)$$

عند اخذ اللوغاريتم إلى الصيغة (2-40) نحصل على الاتي:

$$\log \pi(\beta|x) \propto \log L(\beta) + \log \pi(\beta) \quad \dots (2-41)$$

$$\log \pi(\beta|x) \propto \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} \log e^{X_{ij}^T \beta} + \log (1 + e^{X_{ij}^T \beta})^{-1}] - \frac{1}{2} \beta^T \Sigma^{-1} \beta \quad \dots (2-42)$$

$$\log \pi(\beta|x) \propto \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} X_{ij}^T \beta - \log(1 + e^{X_{ij}^T \beta})] - \frac{1}{2} \beta^T \Sigma^{-1} \beta \quad \dots (2-43)$$

لتقدير المعلمات نحتاج إلى (*Maximum A Posterior MAP*) لحساب الاشتقاق الذي يعظم الدالة اللاحقة. ونلاحظ من الصيغة (2-43) انها تتألف من جزأين جزء اول يخص لوغاريتم دالة الإمكان والجزء الآخر يخص لوغاريتم التوزيع الاولي ولتبسيط الاشتقاق سيتم تجزئته كالاتي:

1-اشتقاق الجزء الأول الذي يتضمن دالة الإمكان

$$\log L(\beta) = \sum_{i=1}^n \sum_{j=1}^J Y_{ij} X_{ij}^T \beta - \log (1 + e^{X_{ij}^T \beta})$$

$$\frac{\partial \log L(\beta)}{\partial (\beta)} = \sum_{i=1}^n \sum_{j=1}^J \left[Y_{ij} X_{ij}^T - \frac{e^{X_{ij}^T \beta}}{1 + e^{X_{ij}^T \beta}} * X_{ij}^T \right] \quad \dots (2-44)$$

$$\varphi(x_{ij}) = \frac{e^{X_{ij}^T \beta}}{1 + e^{X_{ij}^T \beta}} \quad \text{بما ان}$$

عليه يتم تبسيط الصيغة (2-44) لتصبح كالاتي:

$$\frac{\partial \log L(\beta)}{\partial(\beta)} = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] X_{ij} \quad \dots (2-45)$$

2- اشتقاق الجزء الخاص بالتوزيع الاولي وكالاتي:

$$\frac{\partial \log \pi(\beta)}{\partial(\beta)} = \left\{ -\frac{1}{2} \beta^T \Sigma^{-1} \beta \right\} = -\Sigma^{-1} \beta \quad \dots (2-46)$$

وعليه صيغة المشتقة للوغاريتم اللاحق تصبح كالاتي:

$$\nabla_{\beta} \log \pi(\beta|x) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] X_{ij} - \Sigma^{-1} \beta \quad \dots (2-47)$$

إذ ان ∇ يمثل متجه المشتقات الجزئية للوغاريتم التوزيع اللاحق.

ولأجل بلوغ الدالة إلى نهايتها العظمى نساوي المشتقة الجزئية إلى الصفر لنحصل على مقدر الـ MAP وكالاتي:

$$\nabla_{\beta} \log \pi(\beta|x) = \sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] X_{ij} - \Sigma^{-1} \beta = 0 \quad \dots (2-48)$$

نلاحظ من الصيغة المذكورة أنفاً صعوبة التعامل بوضعها الحالي لعدم إمكانية الوصول إلى صيغ تقدير محددة لذا يتم استعمال طرائق تقدير تكرارية وهي طريقة نيوتن رافسون لذا نحتاج إلى اشتقاق مصفوفة هيس.

1- اشتقاق الجزء الخاص بدالة الإمكان

$$\sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})] X_{ij}$$

$$\frac{\partial^2 \log L(\beta)}{\partial(\beta) \partial(\beta^T)} = -\varphi(x_{ij})(1 - \varphi(x_{ij})) X_{ij} X_{ij}^T \quad \dots (2-49)$$

2- اشتقاق الجزء الخاص بالتوزيع الاولي:

$$\frac{\partial^2 \log \pi(\beta)}{\partial(\beta) \partial(\beta^T)} = -\Sigma^{-1} \beta = -\Sigma^{-1} \quad \dots (2-50)$$

اذن مصفوفة هيسين الكلية تكون كالاتي:

$$H(\beta_r) = -\sum_{j=1}^J \sum_{i=1}^n \varphi(x_{ij})(1 - \varphi(x_{ij})) X_{ij} X_{ij}^T - \Sigma^{-1} \quad \dots (2-51)$$

وباتباع خطوات خوارزمية نيوتن رافسون المعرفة في المصدر (19) يتم كتابة المقدر التكراري بطريقة MAP وكالاتي:

$$\beta_{MAP(r+1)} = \beta_r - [H(\beta_r)]^{-1} \nabla_{\beta} \log \pi(\beta_r | x) \quad \dots (2 - 52)$$

$$\begin{aligned} \beta_{MAP(r+1)} = \beta_r &+ \left[\sum_{j=1}^J \sum_{i=1}^n \varphi(x_{ij})^r (1 - \varphi(x_{ij}))^r X_{ij} X_{ij}^T \right. \\ &+ \Sigma^{-1}]^{-1} \left[\sum_{i=1}^n \sum_{j=1}^J [Y_{ij} - \varphi(x_{ij})^r] X_{ij} \right. \\ &\left. - \Sigma^{-1} \beta_r \right] \quad \dots (2 - 53) \end{aligned}$$

2.7 توظيف مقدر الانتروبي مع الاسلوب البيزي بطريقة MAP:

2-7-1 مفهوم الانتروبي العظمى The Concept of Maximum Entropy

تُعد طريقة الانتروبي العظمى إحدى الطرائق الحصينة (Robust Methods)، إذ لا تعتمد على افتراض توزيع محدد للأخطاء العشوائية في الانموذج. وتمتاز هذه الطريقة بإمكانية إعادة نمذجة كل من المعلمات والأخطاء العشوائية بدلالة الاحتمال، وكتابتها على شكل تركيبة محدبة (Convex Combination) تتطلب وجود نقاط دعم (Support Points) مع احتمالات مقابلة لها، يكون عددها مساوياً لعدد نقاط الدعم عادةً ما يتراوح عدد نقاط الدعم بين 2 إلى 7، إلا أن اختيار ثلاث نقاط دعم يعد شائعاً وفعالاً في التطبيقات العملية. وبناءً على ذلك، سيتم اعتماد ثلاث نقاط دعم $M=3$ كما في الصيغة الآتية: [4]

$$\beta_k = [d_{1k} d_{2k} d_{3k}] \begin{bmatrix} p_{1k} \\ p_{2k} \\ p_{3k} \end{bmatrix} = d_k p_k \quad \dots (2 - 54)$$

d_k هو متجه نقاط الدعم

p_k هو متجه الاحتمالات المقابلة لها

2-7-2 مقدر الانتروبي العظمى للمعلمات

Maximum Entropy Estimator for the Parameters

يتطلب توظيف مقدر الانتروبي مع الاسلوب البيزي باستعمال طريقة MAP الحصول أولاً على مقدر الانتروبي العظمى للمعلمات، ثم اعتماد القيمة التقديرية الناتجة عنه كقيمة أولية (*Initial Value*) ضمن خوارزمية التقدير البيزي بطريقة MAP وعن طريق مراجعة بعض الدراسات السابقة [26]، وجد انها اعتمدت على تثبيت الأخطاء العشوائية في أثناء عملية التقدير، في حين تركز هذه الرسالة على تقدير المعلمات فقط بهدف تحسين دقة الانموذج وكفاءة التقدير.

ويُكتب تقدير معلمات انموذج الانحدار اللوجستي بطريقة GME كما يأتي [26] :

$$\hat{\beta}_k^{GME} = \sum_{q=1}^Q d_{qk} \hat{P}_{qk} , \quad \dots (2 - 55)$$

$$k = 0,1,2, \dots K$$

ويمكن إعادة كتابتها بصيغة المصفوفات:

$$\beta = DP \quad \dots (2 - 56)$$

إذ تُمثل \hat{P}_{qk} الاحتمالات التقديرية المقابلة لنقاط الدعم d_q والخاصة بالمعلمة β_k وسيتم اعتماد هذا التقدير كمقدر أولي لطريقة MAP المبينة في الصيغة (2-52)، وذلك بهدف تحسين كفاءة التقدير، إذ إن الهدف الرئيس من تحليل الانحدار هو الحصول على مقدرات كفؤة تُمكن الباحث من التنبؤ.

تعتمد طريقة GME على تعظيم (دالة شانون) للانتروبي كدالة هدف تحت مجموعة من القيود وكما يأتي:

$$H(p_{qk}) = MAX[- \sum_{k=1}^k \sum_{q=1}^Q p_{qk} \ln p_{qk}] \quad \dots (2 - 57)$$

القيود الأول قيد تساوي مجموع الاحتمالات التي تخص المعلمات إلى 1 وكالاتي:

$$\sum_{q=1}^Q p_{qk} = 1 \quad , \quad k = 1, 2, \dots, K \quad , \quad p_{qk} \geq 0 \quad \forall qk$$

القيد الثاني هو قيد الاتساق للنموذج والذي يركز على المعلمات فقط وكالاتي

$$Y = X\beta \rightarrow Y = X(DP) \quad \dots (2 - 58)$$

يمكن توضيح ابعاد الانموذج وكالاتي:

$$Y \text{ هو } n \times 1$$

$$X \text{ هو } n \times (k + 1)$$

$$\beta \text{ هو } (k + 1) \times 1$$

D تمثل مصفوفة نقاط دعم، و p تمثل مصفوفة الاحتمالات.

وعليه يتم كتابة معادلة لاكرانج بدمج دالة الهدف مع القيود باستعمال مضاعفات لاكرانج لنحصل على:

$$l = - \underbrace{\sum_{k=1}^K \sum_{q=1}^Q p_{qk} \ln(p_{qk})}_{\dots} + \underbrace{\lambda' [Y - X(DP)] + \sum_{k=1}^K \alpha_k [\sum_{q=1}^Q p_{qk} - 1]}_{\dots}$$

$$\dots (2 - 59)$$

وبأخذ المشتقة بالنسبة ل p_{qk} ومساواتها إلى الصفر تصبح كالاتي:

$$\frac{\partial l}{\partial p_{qk}} = -\ln p_{qk} - 1 - \lambda' X_k d_{qk} + \alpha_k = 0$$

$$-\ln p_{qk} = 1 - \alpha_k + \lambda' X_k d_{qk} \quad \dots (2 - 60)$$

وبضرب طرفي المعادلة في سالب واحد للصيغة أعلاه نحصل على:

$$\ln p_{qk} = -(1 - \alpha_k + \lambda' X_k d_{qk}) \quad \dots (2 - 61)$$

وبأخذ e لطرفي المعادلة للصيغة المذكورة آنفاً نحصل على:

$$p_{qk} = e^{-(1 - \alpha_k + \lambda' X_k d_{qk})} \quad \dots (2 - 62)$$

$$p_{qk} = e^{\alpha_k - 1} * e^{-\lambda' X_k d_{qk}} \quad \dots (2 - 63)$$

وباستعمال القيد $\sum_{q=1}^Q p_{qk} = 1$ في العلاقة (2 - 63) نحصل على:

$$\sum_{q=1}^Q e^{\alpha_k - 1} * e^{-\lambda' X_k d_{qk}} = 1$$

$$e^{\alpha_k-1} \sum_{q=1}^Q e^{-\lambda' X_k d_{qk}} = 1$$

$$e^{\alpha_k-1} = \frac{1}{\sum_{q=1}^Q e^{-\lambda' X_k d_{qk}}}$$

نعوض e^{α_k-1} في الصيغة (63 - 2) وبعد ترتيب المعادلة نحصل على:

$$\hat{p}_{qk} = \frac{e^{-\lambda' X_k d_{qk}}}{\sum_{q=1}^Q e^{-\lambda' X_k d_{qk}}} \quad \dots (2 - 64)$$

وإذا افترضنا للتبسيط ان متجه مضاعف لاكرانج يمكن إعادة تمثيله بمعلمات الانحدار أي ان $-\lambda' = \beta_k$ لتصبح الصيغة أعلاه كالآتي:

$$\hat{p}_{qk} = \frac{e^{\beta_k X_k d_{qk}}}{\sum_{q=1}^Q e^{\beta_k X_k d_{qk}}} \quad \dots (2 - 65)$$

ومن ثمَّ فإن استعمال قيم \hat{p}_{kq} في الصيغة (2-55) ينتج تقديراً اولياً لمعاملات الانموذج بطريقة GME ويوظف هذا التقدير ضمن صيغة التقدير البيزي بطريقة MAP والتي يمكن إعادة كتابتها:

$$\beta_k^{GME} = \sum_{q=1}^Q \hat{P}_{qk} d_{qk}$$

في هذه الرسالة يُقترح توظيف مقدر الانتروبي العظمى كقيمة أولية في BM، كما في الصيغة (2-52) والتي يتم إعادة كتابتها وكالآتي:

$$\beta_{MAP(r+1)} = \beta_r - [H(\beta_r)]^{-1} \nabla_{\beta} \log \pi(\beta_r | x)$$

نلاحظ أن التقدير في (2-52) باستعمال طريقة نيوتن-رافسون، يمثل تحسناً لمقدر بيز إذ يُعتمد في كل مرة على القيمة السابقة للمعلمة. واستعمال مقدر GME في هذه العملية كقيمة أولية يعد مساهمة بحثية جديدة من لدن الباحثة تهدف إلى تحسين دقة التقدير البيزي وتقديم بديل حصين وفعال للتقدير التقليدي مثل MLE .

2.7.3: خطوات خوارزمية نيوتن-رافسون لتقدير المعلمات بطريقة توظيف مقدر الانتروبي في التقدير البيزي بطريقة *MAP* وكالاتي:

1-الحصول على مقدر الانتروبي من الصيغة (2 – 55)

$$\beta_k^{GME} = \sum_{q=1}^Q \hat{P}_{qk} d_{qk}$$

2-كتابة صيغة *MAP* باستعمال خوارزمية نيوتن-رافسون (2 – 52)

$$\beta_{MAP(r+1)} = \beta_r - [H(\beta_r)]^{-1} \nabla_{\beta} \log \pi(\beta_r | x)$$

3-اعتماد مقدر الانتروبي المعرف في الخطوة الأولى كقيمة أولية في صيغة *MAP* المعرفة في الصيغة (2 – 55).

4-استخراج مقدر *MAP* بعد توظيف مقدر الانتروبي فيه.

5-إعادة تكرار الخطوات من 1 إلى 4 إذا لم يتحقق الشرط $\beta_{MAP(r+1)} - \beta_r < \varepsilon$

2-8: معايير تقييم أداء طرائق التقدير في النماذج اللوجستية

هناك عدة مقاييس تُستخدم لتقييم جودة وأداء طرائق التقدير في النماذج اللوجستية، وفيما يأتي أهمها:

1-متوسط مربعات الخطأ التكاملية:

(Integrated Mean Squared Error – IMSE)

يُعدّ متوسط مربعات الخطأ التكاملية (IMSE) من المقاييس المهمة في تقييم دقة النماذج الإحصائية، ولا سيما في النماذج الاحتمالية كالنماذج اللوجستية المختلطة. ويُستعمل هذا المعيار لقياس مدى انحراف القيم الاحتمالية المقدّرة عن القيم الحقيقية المتوقعة عبر النطاق الكامل لقيم المتغيرات المستقلة، أي أنه يعبر عن متوسط الخطأ التربيعي للتقديرات الاحتمالية على امتداد المجال بدلاً من عند نقطة محددة فقط.

ويُعرّف متوسط مربعات الخطأ التكاملية كالاتي: [22]

$$IMSE(\theta) = E(\hat{\theta}(x) - \theta(x))^2 f(x) dx \quad \dots (2 - 66)$$

2- اختبار هومسر-لامشو (Hosmer-Lemeshow)

هو اختبار لجودة ملائمة نموذج الانحدار اللوجستي، ويهدف إلى التحقق من مدى توافق الاحتمالات المتوقعة من النموذج مع النتائج المشاهدة فعليًا. تعتمد فكرته على ترتيب المشاهدات بحسب الاحتمال المتوقع \hat{p}_i ثم تقسيمها إلى g مجموعات (غالبًا 10 مجموعات متساوية تقريبًا)، وبعد ذلك تتم مقارنة عدد/مجموع القيم المشاهدة للحالات $y=1$ مع القيم المتوقعة داخل كل مجموعة. يُحسب إحصاء الاختبار على صورة كاي-تربيع كما يأتي: [18]

$$\chi^2_{HL} = \sum_{k=1}^g \left[\frac{(O_{1k} - E_{1k})^2}{E_{1k}} + \frac{(O_{0k} - E_{0k})^2}{E_{0k}} \right],$$

حيث O_{1k} هو عدد الحالات المشاهدة ذات $y=1$ في المجموعة k ، $E_{1k} = \sum \hat{p}_i$ هو العدد المتوقع لها، وكذلك O_{0k} و $E_{0k} = \sum (\hat{p}_i - 1)$ للحالات ذات $y=0$. درجات الحرية عادة $df=g-2$. فرضية العدم H_0 تنص على حسن ملائمة النموذج 0 لافروق جوهري بين المشاهد والمتوقع) وعليه فان قيمة $p\text{-value} > 0.05$ تشير الى عدم وجود دليل احصائي على سوء الملاءمة، اي ان النموذج مناسب للبيانات.

3- اختبار نسبة الإمكانان: (Likelihood Ratio Test)

يُعدّ اختبار نسبة الإمكانان من الاختبارات الإحصائية الأساسية لتقييم معنوية النماذج اللوجستية، ولا سيما ضمن إطار النمذجة الهرمية. ويُستعمل للمقارنة بين نموذجين متداخلين (Nested Models) الانموذج المقيد (Restricted Model) او الانموذج الصفري الذي يستبعد بعض المعاملات أو التأثيرات العشوائية والانموذج غير المقيد (Unrestricted Model) أي الانموذج الذي يتضمن جميع المتغيرات التوضيحية أو التأثيرات العشوائية المراد اختبارها، وتُحسب قيمة إحصاء الاختبار كما يأتي: [19]

$$D_{01} = -2 \log \left[\frac{L_0}{L_1} \right] \quad \dots (2 - 68)$$

وتُقارن قيمة (D) بتوزيع كاي-تربيع (χ^2) لتحديد قيمة الدلالة الإحصائية (sig) لتحديد معنوية الفرق بين النموذجين.

الفصل الثالث

الجانب التجريبي

3.1 تمهيد (Preface) :

تم في هذا الفصل اجراء تحليلًا تجريبيًا مقارنًا لثلاث طرائق لتقدير المعلمات في النماذج الإحصائية، وهي طريقة (MLE)، وطريقة (BM)، وطريقة (BME)، وذلك بهدف تقييم أدائها ضمن سياقات مختلفة تتعلق بحجم العينة، وعدد المجموعات، وعدد المتغيرات المستقلة في مستويات البنية المختلفة. وقد تم تنفيذ سلسلة من التجارب العددية خطوات محاكاة مونت- كارلو والتي تتضمن تباينات متعددة في البنية التجريبية، لتحديد الظروف التي تبرز فيها أفضلية كل طريقة على حدة. وشمل التحليل عدد مرات التفوق لكل طريقة ضمن جداول مفصلة. كما تم تفسير النتائج في ضوء الاعتبارات النظرية لكل طريقة، مع التركيز على مدى استقرارها ودقتها في ظل ظروف البيانات ذات الخصائص المختلفة. يهدف هذا الفصل إلى استخلاص مؤشرات منهجية تساعد في توجيه الباحثين لاختيار طريقة التقدير الأنسب على وفق طبيعة البيانات وتعقيد الانموذج الإحصائي المستعمل.

3.2 مفهوم المحاكاة (Simulation Concept)

تُعد المحاكاة (Simulation) إحدى الأدوات العلمية المتقدمة التي تُستعمل لنمذجة الواقع أو أحد عناصره بشكل رياضي أو منطقي من أجل تحليل سلوك نظام من دون الحاجة إلى التدخل الفعلي في ذلك النظام، ما يوفر الجهد، الوقت، والتكلفة، ويقال من المخاطر المحتملة عند اختبار التغيرات في البيئات الحقيقية. ويعتمد مبدأ المحاكاة على بناء انموذج يعكس خصائص النظام الواقعي، ثم تشغيل هذا الانموذج افتراضياً ضمن مجموعة من السيناريوهات المدروسة لدراسة مخرجاته، ما يمكن الباحث أو صانع القرار من التقدير بالأداء المستقبلي للنظام، أو تقييم مدى تأثير التعديلات المقترحة عليه. [25]

تنقسم المحاكاة إلى عدة أنواع وفقاً لطبيعة النظام المدروس وأسلوب التمثيل المستعمل. من أبرز أنواع المحاكاة هي المحاكاة المنفصلة الأحداث (Discrete Event Simulation)، والتي تُستعمل عندما يكون النظام يتغير فقط في لحظات معينة من الزمن نتيجة لحدوث أحداث منفصلة، مثل محاكاة طابور الانتظار في المستشفى أو المطار؛ والمحاكاة المستمرة (Continuous Simulation)، التي تُعنى بدراسة الأنظمة التي تتغير حالتها بشكل مستمر عبر الزمن كما في النماذج الفيزيائية أو الكيميائية (مثلاً، تغيير درجة الحرارة أو الضغط). كذلك، هناك المحاكاة الحتمية (Deterministic Simulation)، التي لا تتضمن أي عنصر من عناصر العشوائية وتنتج دوماً النتائج نفسها عند تكرارها، بينما تعتمد المحاكاة الاحتمالية (Stochastic Simulation)

(Simulation) على إدخال متغيرات عشوائية بحيث تعكس التباين الطبيعي أو عدم التيقن الموجود في الواقع، وتُستعمل بشكل واسع في النماذج الاقتصادية والطبية والبيئية [24].

أما من الناحية الإحصائية، فالمحاكاة تُعد حجر زاوية في الكثير من التحليلات الحديثة، إذ تُستعمل في عدة مجالات تشمل تقدير القيم الإحصائية عندما يصعب الحصول عليها بالأساليب التحليلية، واختبار الفرضيات، وحساب فترات الثقة، وتحليل الأداء تحت ظروف افتراضية. ومن أبرز أدوات المحاكاة الإحصائية ما يُعرف باسم محاكاة مونت كارلو (Monte Carlo Simulation)، وهي تقنية تعتمد على توليد أعداد كبيرة من القيم العشوائية المستمدة من توزيعات معينة، ثم استعمال هذه القيم لمحاكاة مخرجات نظام أو عملية معينة عدة مرات بهدف استكشاف سلوكها العام أو استقرارها. تُستعمل هذه التقنية في تقييم المخاطر، التقدير المالي، النمذجة البيئية، وكذلك في تصميم التجارب الإحصائية وتقييم قوة الاختبارات. كما تُستعمل المحاكاة في الإحصاء لتحليل تأثير الأخطاء العشوائية أو الانحرافات المنهجية، وتُعد وسيلة فعالة لتقدير التوزيعات العينية، لاسيما عندما تكون العينة صغيرة أو لا تتبع التوزيع الطبيعي. وفضلاً عن ذلك، أصبحت المحاكاة ذات أهمية بالغة في تطبيقات الذكاء الاصطناعي، تعلم الآلة، والنماذج الاحتمالية المعقدة مثل النمذجة البيزية، إذ تُمثل الطريقة الوحيدة الممكنة لتوليد البيانات أو فحص أداء النماذج في بيئة افتراضية قابلة للضبط والتحكم [30].

3.3 مراحل وخطوات المحاكاة (Simulation stages and steps)

تم إجراء تجارب محاكاة مونت كارلو لبيان افضلية طرائق التقدير عن طريق بناء نموذج انحدار لوجستي بمستويين وتم تصميم النموذج بشكل مرن يسمح بزيادة أو تقليل عدد المجموعات وعدد المتغيرات التوضيحية على كلا المستويين، لتقييم أثر عدد المتغيرات في دقة التقدير على وفق الخطوات الآتية:

3.3.1 تحديد القيم الافتراضية لمعاملات النماذج المدروسة في تجارب المحاكاة وتعد هذه الخطوة من الخطوات المهمة التي تعتمد عليها تجارب المحاكاة اللاحقة وفقاً لما يأتي:

- تحديد عدد المجموعات G إذ تم اختيار مجموعات مختلفة تم عندها بناء نماذج تجارب المحاكاة وهي:

$$G=2, 3, 5$$

- تحديد حجم العينات داخل كل مجموعة n اذ تم اختيار أربع احجام عينات (صغيرة – متوسطة – كبيرة) وكالاتي:

$$n=25, 50, 75, 100$$

- تحديد عدد المتغيرات التوضيحية في المستوى الأول Level1 Predictors = K
- عدد المتغيرات التوضيحية في المستوى الثاني Level2 Predictors = M
- اختيار القيم الافتراضية لثوابت النموذج الانحدار وكالاتي:

3.3.2 توليد التأثيرات الثابتة (Fixed effects)

- γ_{00} وهي المعلمة الأساسية التي تمثل الحد الثابت (intercept) في الانموذج ، وهي تمثل القيمة المتوقعة لـ log-odds عندما تكون جميع المتغيرات الأخرى (X) و (W) مساوية إلى الصفر والتي تمثل التأثيرات الثابتة (Fixed effects) وتم توليدها توافقاً مع كل انموذج من توزيع طبيعي قياسي وكالاتي:

$$\text{Gamma}_{00} = \text{randn}();$$

- γ_{k0} هو متجه معاملات الميل (slopes) المرتبطة بالمتغيرات على المستوى الأول (Level-1 predictors) وهي تمثل التأثيرات الثابتة لكل متغير على مستوى الفرد أو الوحدة. وتم توليده من توزيع طبيعي قياسي على وفق الاتي:

$$\text{Gamma}_{k0} = \text{randn}(\text{nLevel1Predictors}, 1);$$

- γ_{0m} هو متجه معاملات يمثل تأثير متغيرات المستوى الثاني (Level-2 predictors) على الحد الثابت (intercept) في الانموذج. أي أنه يوضح كيف تؤثر خصائص المجموعة (وليس الفرد) في نقطة البداية (baseline log-odds) للانحدار. وتم توليده من توزيع طبيعي قياسي على وفق الاتي:

$$\text{Gamma}_{0m} = \text{randn}(1, \text{nLevel2Predictors});$$

- γ_{km} وهو من أهم الأجزاء في الانموذج؛ لأنه يُعبّر عن تفاعلات المستويين الأول والثاني (Cross-Level Interactions) , وهو مصفوفة معاملات تمثل التفاعلات بين متغيرات

المستوى الأول (Level-1) ومتغيرات المستوى الثاني (Level-2) , بمعنى ان كل خلية فيها توضّح كيف يؤثر التفاعل بين متغير فردي ومتغير جماعي في مخرجات الانموذج , وتم توليده من توزيع طبيعي قياسي على وفق الاتي:

```
Gamma_km = randn(nLevel1Predictors,
nLevel2Predictors);
```

3.3.3 اختيار تباينات التأثيرات العشوائية (Random effects)

- للحد الثابت للمجموعة z ، بالقيمة 0.5

```
sigmaU_intercept=0.5;
```

- عبر المجموعات، لكل مجموعة بالقيمة 0.4

```
sigmaU_slopes=0.4;
```

==تم اختيار قيم تباينات التأثيرات العشوائية في المحاكاة، وهي 0.5 للحد الثابت و0.4 للميل، بناءً على مبدأ تمثيل سيناريو معتدل يُظهر التباين بين المجموعات من دون أن يكون مفرطاً، ما يسمح باختبار أداء طرائق التقدير المختلفة تحت ظروف متوازنة. فالقيم الصغيرة جداً قد تؤدي إلى إهمال التأثيرات العشوائية، بينما القيم الكبيرة قد تسبب عدم استقرار عددي في الانموذج ، لذلك فإن اختيار هذه القيم يحقق توازناً بين الواقعية والاستقرار الرياضي. كما أن هذه القيم شائعة في الأدبيات الإحصائية الخاصة بنماذج الانحدار متعدد المستويات، وتعتمد غالباً في دراسات المحاكاة المنشورة التي تهدف إلى تقييم دقة التقدير والتحقق من حساسية النماذج لتغير التباين [5].

3.3.4 توليد بيانات المتغيرات (Variables generation)

في هذه المرحلة يتم على وفق الخطوات الاتية:

- 1- حساب الحجم الكلي للعَيِّنات (المشاهدات) عن طريق ضرب عدد المجموعات \times عدد الأفراد داخل كل مجموعة كالآتي:

```
totalSamples = numGroups * samplesPerGroup;
```

- 2- إنشاء متجه group يحدد إلى أي مجموعة ينتمي كل فرد كالآتي:

```
group = repelem(1:numGroups, samplesPerGroup)';
```

3- توليد بيانات المتغيرات التوضيحية على المستوى الثاني (W) بحيث ان كل صف يمثل مجموعة، وكل عمود يمثل أحد متغيرات المستوى الثاني كالآتي:

```
W = randn(numGroups, nLevel2Predictors);
```

4- توليد بيانات المتغيرات التوضيحية على المستوى الاول (X) بحيث ان كل صف يمثل فرداً ، وكل عمود يمثل احد متغيرات المستوى الاول كالآتي:

```
X = randn(totalSamples, nLevel1Predictors);
```

5- توليد القيم العشوائية للحد الثابت العشوائية لكل مجموعة موزعة طبيعياً بمتوسط صفر وانحراف معياري sigmaU_intercept السابق كالآتي:

```
U0 = normrnd(0, sigmaU_intercept, numGroups, 1);
```

6- توليد القيم العشوائية للميل العشوائي لكل مجموعة ولكل متغير على المستوى الأول كالآتي:

```
Uk = normrnd(0, sigmaU_slopes, numGroups, nLevel1Predictors);
```

7- تهيئة متجه y (المتغير التابع) لحفظ نتائج المحاكاة كالآتي:

```
y = zeros(totalSamples, 1);
```

8- انشاء حلقة تكرارية (for-loop) لـ كل فرد من 1 إلى (total Samples) يتم تنفيذ كما يأتي:

- تحديد رقم المجموعة التي ينتمي إليها الفرد الحالي كالآتي:

```
j = group(i);
```

- استخراج متجه متغيرات الفرد X_i واستخراج خصائص المجموعة W_j كالآتي:

```
x_i = X (i, :)' ;
```

```
w_j = W (j, :)' ;
```

- حساب الجزء الثابت (Fixed Effects) كالآتي:

```
fixed = Gamma_00 + dot (Gamma_k0, x_i) + dot (Gamma_0m, w_j);
```

- إضافة التأثيرات التفاعلية (Cross-Level) كالآتي:

```

for k = 1: nLevel1Predictors
    for m = 1: nLevel2Predictors
        fixed = fixed + Gamma_km(k, m) * x_i(k) *
w_j(m);
    end
end
end

```

- جمع كل التأثيرات التفاعلية بين X و W، كالاتي:

Gamma_km(k, m)

- حساب الجزء العشوائي (Random Effects) كالاتي:

random = U0(j) + Uk(j, :) * x_i;

- حساب logit والاحتمال كالاتي:

linpred = fixed + random;

p = 1 / (1 + exp(-linpred));

- توليد المتغير التابع ثنائي القيمة (Y) بناء على قيمة الاحتمال P من توزيع ثنائي الحدين (Bernoulli trial) كالاتي:

y(i) = binornd(1, p);

3.3.5 مقارنة النتائج:

تم اعتماد معايير أساسية لتقييم أداء طرائق التقدير في انموذج الانحدار اللوجستي متعدد المستويات، وهي متوسط مربعات الخطأ التكاملية (IMSE) المبين في الفقرة (2-8) من الفصل الثاني الذي يُستعمل لقياس دقة التقدير الاحتمالي عبر النطاق الكامل لقيم المتغيرات، كذلك تم استعمال قيمة الدلالة الإحصائية (Sig.) الناتجة عن اختبار نسبة الإمكانات (Likelihood Ratio Test) لتحديد ما إذا كان الانموذج المدروس يحقق دلالة معنوية مقارنةً بالأنموذج الفارغ (Null model).

3.3.6 تحليل تجارب المحاكاة:

تم اعتماد طرائق التقدير الآتية في تقدير انموذج الانحدار اللوجستي بمستويين وهي:

طريقة الإمكان الأعظم (MLE) المبينة في الفقرة (2-5-1) من الفصل الثاني

طريقة بيز- ماب (BM) المبينة في الفقرة (2-6) من الفصل الثاني

طريقة بيز- ماب انتروبي (BME) المبينة في الفقرة (2-7) من الفصل الثاني وكانت النتائج كما في الجداول الآتية :

المثال الاول (Example one)

تم افتراض انموذج انحدار لوجستي ثنائي المستوى يُستعمل لتقدير احتمال حدوث نتيجة ثنائية \hat{Y}_{ij}^* للفرد i ضمن المجموعة j ، وذلك عن طريق دمج متغيرات تفسيرية على مستويين وبحسب المدخلات الآتية:

عدد المجموعات $G=2$

عدد متغيرات المستوى الاول $K=2$

عدد متغيرات المستوى الثاني $M=2$

على وفق انموذج التأثيرات الثابتة (Fixed Effects Model) الآتي:

$$\begin{aligned} \hat{Y}_{ij}^* &= \gamma_{00} + \sum_{k=1}^K \gamma_{k0} X_{kij} + \sum_{m=1}^M \gamma_{0m} W_{mj} + \sum_{m=1}^M \sum_{k=1}^K \gamma_{km} W_{mj} X_{kij} + \sum_{k=1}^K u_{kj} X_{kij} \\ &= \gamma_{00} + \sum_{k=1}^2 \gamma_{k0} X_{kij} + \sum_{m=1}^2 \gamma_{0m} W_{mj} + \sum_{m=1}^2 \sum_{k=1}^2 \gamma_{km} W_{mj} X_{kij} \\ &= \gamma_{00} + \gamma_{10} X_{1ij} + \gamma_{20} X_{2ij} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + \gamma_{11} W_{1j} X_{1ij} \\ &\quad \gamma_{21} W_{1j} X_{2ij} + \gamma_{12} W_{2j} X_{1ij} + \gamma_{22} W_{2j} X_{2ij} \quad \dots (3-1) \end{aligned}$$

عند حجم عينة $n=25$:

جدول (1-3) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الاول عند حجم عينة $n=25$

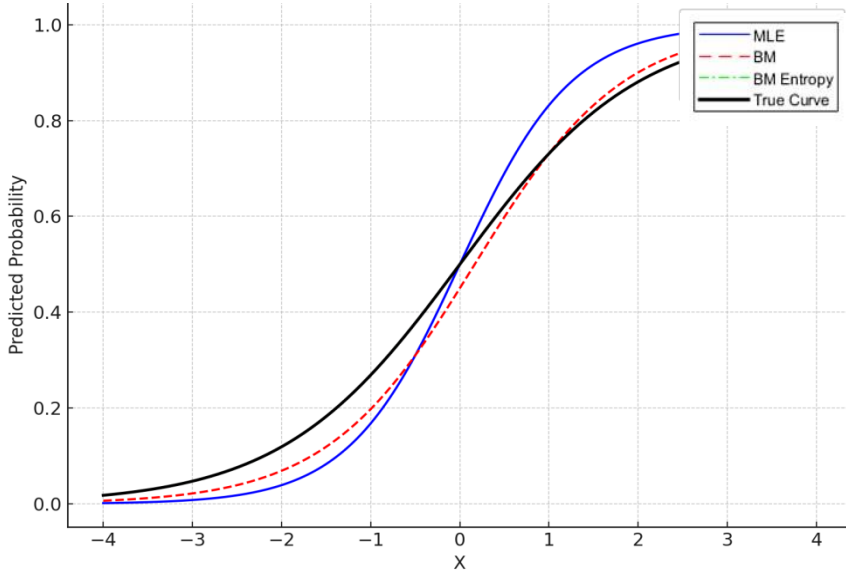
Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	0.500	0.114	0.392	0.663	0.021	0.011	0.001
γ_{10}	-1.563	-0.425	-0.978	-1.121	0.011	0.010	0.001
γ_{20}	-0.485	-0.181	-0.259	-0.544	0.022	0.012	0.002
γ_{01}	1.295	0.482	1.239	1.561	0.042	0.022	0.012
γ_{02}	-1.207	-0.916	-1.031	-1.303	0.004	0.003	0.002
γ_{11}	0.703	0.254	0.952	0.825	0.025	0.018	0.012
γ_{21}	-1.453	-0.964	-0.896	-1.247	0.013	0.011	0.011
γ_{12}	-0.667	-0.426	-1.075	-0.945	0.044	0.023	0.011
γ_{22}	0.504	0.133	0.318	0.638	0.089	0.042	0.012

يتضح من الجدول (1-3) ان طريقة (BME) هي الأكثر دقة والأفضل أداءً من بين طرائق التقدير الباقية (MLE) ، (BM) ، تليها طريقة BM ومن ثم طريقة MLE من حيث القرب من القيم الحقيقية بحيث سجلت أدنى انحراف عن القيمة الأصلية في معظم المعاملات ما يعكس دقة واضحة في استرجاع البنية الحقيقية للأنموذج . كذلك، من حيث مستوى الدلالة الإحصائية (P-value) ، كانت (BME) أيضاً الفضلى، إذ أعطت أصغر قيم للـ P-values في جميع المعاملات تقريباً، ما يدل على أن التقديرات الناتجة عنها كانت الأكثر معنوية وثقة.

الفصل الثالث الجانبي التجريبي

جدول (2-3) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الاول عند حجم عينه n=25

Y_{ij}^*	\hat{Y}_{ijMLE}^*	\hat{Y}_{ijBM}^*	\hat{Y}_{ijBME}^*	$MSE\hat{Y}_{ijMLE}^*$	$MSE\hat{Y}_{ijBM}^*$	$MSE\hat{Y}_{ijBME}^*$
0.988355	0.846181	0.968281	0.989518	0.020213	0.000403	0.000001
0.987688	0.841655	0.966029	0.988535	0.021326	0.000469	0.000001
0.981451	0.837021	0.963622	0.987462	0.020860	0.000318	0.000036
0.986445	0.832279	0.961052	0.986289	0.023767	0.000645	0.000000
0.983468	0.827427	0.958309	0.985009	0.024349	0.000633	0.000002
0.982146	0.822465	0.955381	0.983610	0.025498	0.000716	0.000002
0.981134	0.817392	0.952258	0.982084	0.026811	0.000834	0.000001
0.979331	0.812207	0.948927	0.980419	0.027930	0.000924	0.000001
0.975633	0.806909	0.945378	0.978602	0.028468	0.000915	0.000009
0.976542	0.801499	0.941598	0.976620	0.030640	0.001221	0.000000
IMSE				0.129715	0.102610	0.097352
				3	2	1
Sig.	0.000	0.000	0.000	Rank		



شكل (1-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=25$.

يتضح من جدول (2-3) والشكل (1-3) وعند حجم عينه $n=25$ ما يأتي:

1- أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تليها طريقة BM , واخيراً جاءت طريقة (MLE)، في المرتبة الأخيرة. اذ نلاحظ ان الاحتمالات المتوقعة كانت متقاربة للاحتمالات الحقيقية عند طريقة (BME) وانها سجلت اقل متوسط مربعات خطأ (MSE) لكل مشاهدة على حدة مقابل قيم اعلى بكثير لطريقتي (MLE , BM) ما يشير إلى دقة عالية جداً لطريقة (BME) في التقدير.

2- ان متوسط مربعات الخطأ التكاملي (IMSE) عند طريقة (BME) سجل ادنى القيم والذي بلغ (0.097352) يأتي بعدها طريقة BM بقيمة بلغت (0.102610) , واخيراً طريقة MLE والذي بلغ (0.129715) ما يعزز أفضلية (BME) في الأداء العام للأنموذج .

3- ان قيمة الدلالة الإحصائية $\text{Sig.} = 0.000$ تشير إلى أن الانموذج معنوي بدرجة عالية.

عند حجم عينة n=50:

جدول (3-3) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الاول عند حجم عينة n=50

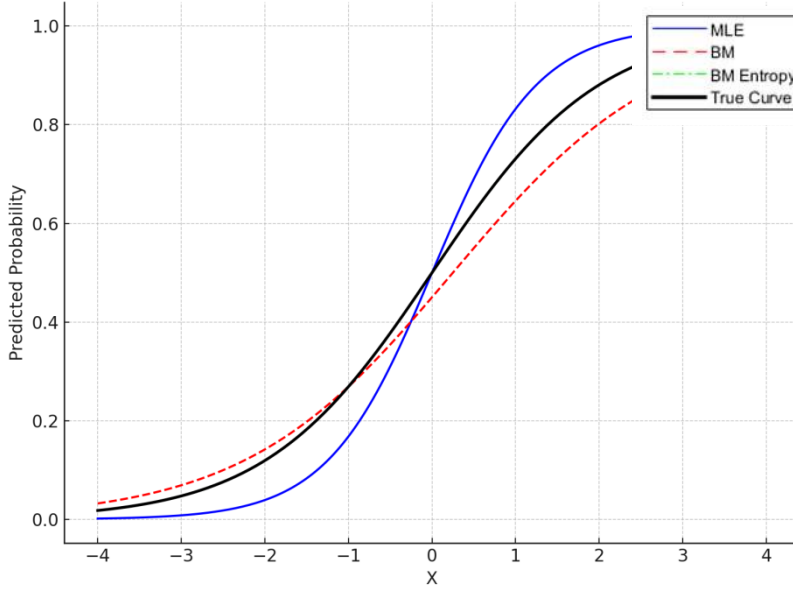
Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
Y ₀₀	0.500	0.628	0.466	0.581	0.000	0.000	0.000
Y ₁₀	-1.563	-1.797	0.360	-0.593	0.000	0.002	0.000
Y ₂₀	-0.485	-0.511	0.290	-0.473	0.000	0.019	0.000
Y ₀₁	1.295	1.241	0.333	1.206	0.000	0.022	0.000
Y ₀₂	-1.207	0.869	0.492	1.181	0.000	0.009	0.000
Y ₁₁	0.703	0.872	1.715	0.722	0.000	0.000	0.000
Y ₂₁	-1.453	-1.415	-0.017	-1.422	0.000	0.036	0.000
Y ₁₂	-0.667	-0.267	-0.062	-0.518	0.000	0.001	0.000
Y ₂₂	0.504	0.580	0.220	0.619	0.000	0.001	0.000

يتضح من الجدول (3-3) أن طريقة (BME) لا تزال تقدم أداءً مميزاً من حيث دقة التقدير وقوة الدلالة. فعند مقارنة القيم المقدرة مع القيم الحقيقية، نلاحظ أن طريقة (BME) كانت الأقرب للقيم الحقيقية في معظم المعاملات مثل γ_{20} ، γ_{11} ، γ_{12} ، γ_{21} ، ما يدل على قدرتها العالية في استرجاع البنية الأصلية للأنموذج. كما أن طريقة (BME) أعطت أصغر قيم للـ P-values في جميع المعاملات (0.000)، ما يعكس أعلى مستويات الدلالة والثقة الإحصائية في نتائجها. وعلى الرغم من أن طريقة (MLE) أظهرت أحياناً تقارباً مع القيم الحقيقية، لاسيما في γ_{01} ، γ_{22} ، فإنها في معاملات أخرى مثل γ_{10} و γ_{02} انحرفت بشكل واضح، وهو ما يحد من موثوقيتها العامة. أما طريقة (BM) فقد أظهرت تفاوتاً كبيراً في التقديرات، وكانت في بعض الحالات مثل (γ_{11} γ_{21}) بعيدة تماماً عن القيم الحقيقية، ما يعكس ضعفاً نسبياً في دقة الاستدلال. وعليه، يمكن القول إن طريقة (BME) تتفوق بوضوح على بقية الطرائق، سواء من حيث القرب من القيم الحقيقية أو من حيث قوة الدلالة الإحصائية، ما يجعلها الأنسب لاعتمادها في مثل هذه النماذج متعددة المستويات عند هذا الحجم من العينة.

الفصل الثالث الجانب التجريبي

جدول (3-4) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الاول عند حجم عينه n=50

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.98023	0.98347	0.98417	0.98111	0.000010	0.000016	0.000001
0.79288	0.79867	0.79120	0.79047	0.000034	0.000003	0.000006
0.18573	0.13230	0.28590	0.18806	0.002855	0.010034	0.000005
0.17880	0.12508	0.29187	0.18069	0.002886	0.012785	0.000004
0.17207	0.11820	0.29792	0.18054	0.002902	0.015838	0.000072
0.16554	0.11165	0.30404	0.17548	0.002904	0.019182	0.000099
0.15921	0.10542	0.31023	0.17196	0.002893	0.022807	0.000136
0.15308	0.09950	0.31648	0.15587	0.002871	0.026700	0.000008
0.14714	0.09387	0.32281	0.14726	0.002838	0.030860	0.000000
0.14140	0.08853	0.32920	0.13614	0.002795	0.035269	0.000028
IMSE				0.002678	0.121353	0.000435
				2	3	1
Sig.	0.006	0.000	0.000	Rank		



شكل (2-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه $n=50$

يتضح من الجدول (3-4) والشكل (2-3)، وعند حجم عينة $n=50$ ، ما يأتي:

- 1- أن طريقة (BME) قدمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
- 2- ان متوسط مربعات الخطأ التكاملي (IMSE)، سجلت عنده طريقة BME أدنى قيمة بلغت (0.000435)، تلتها طريقة MLE بقيمة بلغ (0.002678) واخيرا طريقة BM بقيمة بلغت (0.121353)، وهذا يدل على تفوق BME من حيث الأداء الكلي للأنموذج.

3- تشير قيمة الدلالة الإحصائية (Sig. = 0.000) لجميع الطرائق إلى أن الانموذج معنوي بدرجة عالية، لكن التقديرات الناتجة عن طريقة BME كانت الأكثر دقة ودلالة، ما يجعلها الطريقة الأكثر موثوقية لاعتمادها في النماذج اللوجستية متعددة المستويات عند هذا الحجم من العينة.

عند حجم عينه $n=75$:

جدول (3-5) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الاول عند حجم عينه $n=75$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BMMM E}	P-Value _{MLE}	P-Value _{BM}	P- Value _{BME}
γ_{00}	0.500	0.561	0.282	0.529	0.000	0.022	0.000
γ_{10}	-1.563	-1.685	-0.502	-1.572	0.000	0.037	0.000
γ_{20}	-0.485	-0.502	0.550	-0.530	0.000	0.000	0.000
γ_{01}	1.295	1.418	-0.305	1.353	0.000	0.001	0.000
γ_{02}	-1.207	-1.453	-1.667	-1.224	0.000	0.013	0.000
γ_{11}	0.703	1.083	0.466	0.733	0.000	0.022	0.000
γ_{21}	-1.453	-1.503	-1.941	-1.454	0.000	0.034	0.000
γ_{12}	-0.667	-0.555	-0.155	-0.652	0.000	0.000	0.000
γ_{22}	0.504	0.486	0.293	0.451	0.000	0.001	0.000

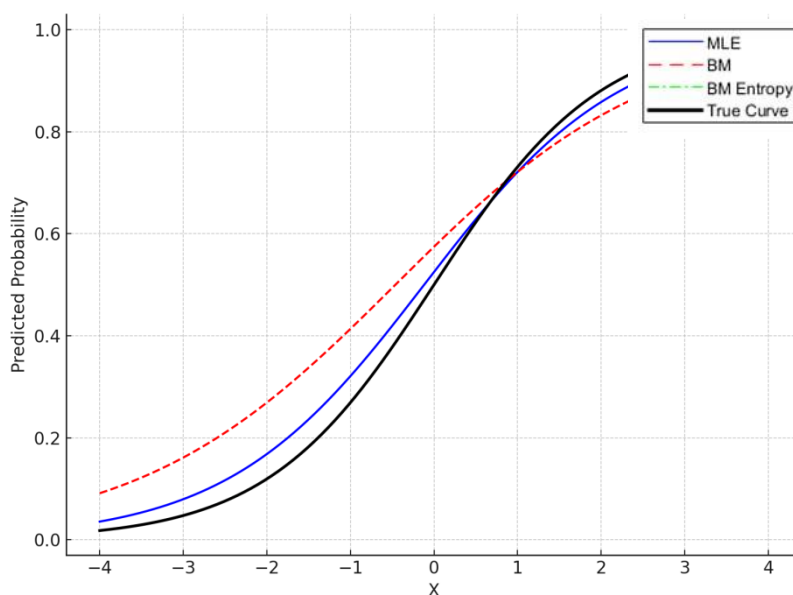
يتضح من الجدول (3-5) أن طريقة (BME) تميزت بتقديم أفضل أداء من حيث دقة التقدير وقوة الدلالة الإحصائية مقارنةً بطرائق التقدير الأخرى. فقد كانت تقديرات (BME) الأقرب إلى القيم الحقيقية في معظم المعاملات، مثل γ_{00} ، γ_{10} ، γ_{20} ، γ_{01} ، γ_{02} ، γ_{11} ، و γ_{21} ، ما يدل على قدرتها العالية على استرجاع الهيكل الحقيقي للأنموذج. كما أن جميع معاملات (BME) جاءت ذات دلالة إحصائية عالية، إذ سجلت $P\text{-value} = 0.000$ في المعاملات كافة، وهو ما يعكس أقصى درجات المعنوية والثقة في نتائج التقدير. في المقابل، أظهرت (MLE) أداءً مقبولاً في بعض المعاملات، مثل γ_{20} و γ_{22} ، إلا أنها انحرقت بشكل واضح في معاملات أخرى مثل γ_{02} و γ_{11} ، وهو ما يضعف من موثوقيتها بشكل عام. أما طريقة (BM) فقد أظهرت انحرافاً كبيراً في عدة معاملات (لاسيما γ_{01} و γ_{10} و γ_{11})، كما سجلت مستويات دلالة ضعيفة نسبياً في بعضها، مثل γ_{10} ($P = 0.037$) و γ_{11} ($P = 0.022$)، ما يشير إلى ضعف في الدقة والمعنوية مقارنة بباقي الطرائق.

الفصل الثالث الجانب التجريبي

جدول (6-3) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للأنموذج الاول عند حجم عينه

n=75

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.141395	0.139998	0.251129	0.150802	0.000002	0.012042	0.000009
0.147140	0.142835	0.256406	0.153977	0.000019	0.011939	0.000005
0.153077	0.145864	0.261832	0.157338	0.000052	0.011828	0.000002
0.159208	0.149097	0.267408	0.160895	0.000102	0.011707	0.000000
0.165537	0.152544	0.273137	0.164658	0.000169	0.011578	0.000000
0.172067	0.156220	0.279020	0.168636	0.000251	0.011439	0.000001
0.178798	0.160136	0.285058	0.172840	0.000348	0.011291	0.000004
0.185734	0.164306	0.291252	0.177280	0.000459	0.011134	0.000007
0.192876	0.168745	0.297603	0.181967	0.000582	0.010968	0.000012
0.200225	0.173466	0.304112	0.186911	0.000716	0.010793	0.000018
IMSE				0.000567	0.111278	0.000078
				2	3	1
Sig.	0.000	0.001	0.000	Rank		



شكل (3-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينه n=75

يتضح من الجدول (6-3) والشكل (3-3)، وعند حجم عينة $n=75$ ، ما يأتي:

- 1- أن طريقة (BME) قدمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
- 2- أما بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت عند (BME) أدنى قيمة بلغ (0.000078)، تلتها طريقة (MLE) بقيمة بلغت (0.000567)، وأخيراً طريقة (BM) بقيمة مرتفعة بلغت (0.111278). وهذا يدل بوضوح على تفوق BE في الأداء الكلي للأنموذج من حيث التقدير.

- 3- تشير قيمة الدلالة الإحصائية (Sig. = 0.000) لجميع الطرائق إلى أن الانموذج معنوي بدرجة عالية، إلا أن التقديرات الناتجة عن (BME) كانت الأكثر دقة وثباتاً، ما يجعلها الطريقة الأكثر موثوقية لاعتمادها في النماذج اللوجستية متعددة المستويات عند هذا الحجم من العينة.

عند حجم عينه $n=100$:

جدول (7-3) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الأول عند حجم عينه $n=100$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	0.500	0.405	0.121	0.620	0.000	0.021	0.000
γ_{10}	-1.563	-1.465	0.424	-1.559	0.000	0.021	0.000
γ_{20}	-0.485	-0.581	-0.016	-0.487	0.000	0.000	0.000
γ_{01}	1.295	1.332	0.533	1.203	0.000	0.000	0.000
γ_{02}	-1.207	-1.292	-0.331	-1.195	0.000	0.011	0.000
γ_{11}	0.703	0.725	0.662	0.710	0.000	0.012	0.000
γ_{21}	-1.453	-1.534	0.771	-1.357	0.000	0.022	0.000
γ_{12}	-0.667	-0.567	0.010	-0.652	0.000	0.000	0.000
γ_{22}	0.504	0.789	0.339	0.564	0.000	0.001	0.000

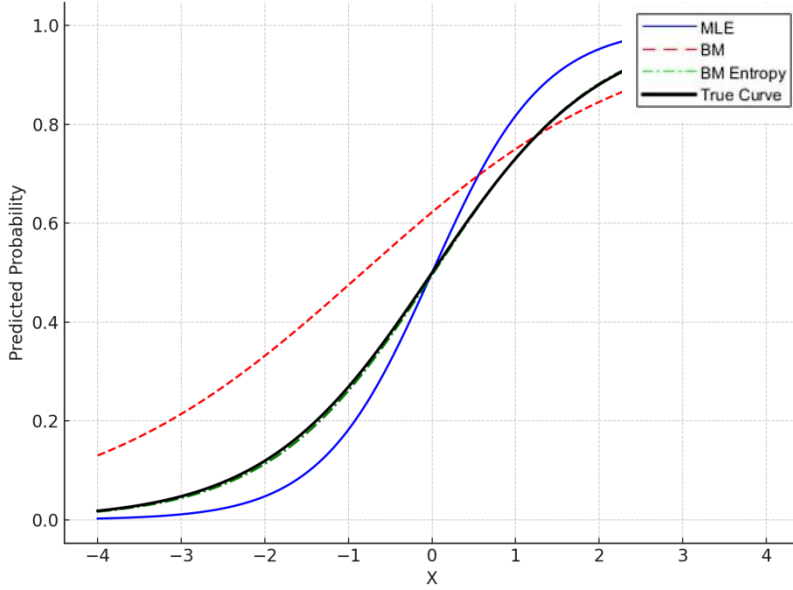
يتضح من الجدول (7-3) أن طريقة (BME) قد تميزت بتقديم أفضل أداء من حيث دقة التقدير وقوة الدلالة الإحصائية مقارنةً بطرائق التقدير الأخرى. فقد كانت تقديرات (BME) هي الأقرب إلى القيم الحقيقية في معظم المعاملات، مثل γ_{00} ، γ_{10} ، γ_{20} ، γ_{01} ، γ_{02} ، γ_{11} ، وهو ما يدل بوضوح على قدرتها العالية في استرجاع الهيكل الحقيقي للأنموذج بدقة وموثوقية. سجلت جميع معاملات

الفصل الثالث الجانب التجريبي

(BME) مستوى دلالة إحصائية مرتفع (P-value = 0.000) ، وهو ما يعكس أقصى درجات المعنوية والثقة في نتائج التقدير، ويؤكد أن الانموذج الناتج عنها مستقر وموثوق في جميع المعاملات. في المقابل، أظهرت طريقة (MLE) أداءً مقبولاً في بعض المعاملات مثل γ_{00} و γ_{02} ، إلا أنها انحرفت بشكل ملحوظ في معاملات أخرى مثل γ_{22} ، وهو ما يؤدي إلى تذبذب في موثوقيتها العامة عند مقارنة نتائجها مع القيم الحقيقية. أما طريقة (BM) فقد أظهرت انحرافات كبيرة في عدة معاملات، لاسيما في γ_{01} و γ_{10} و γ_{21} ، إذ ابتعدت القيم المقدرة كثيراً عن القيم الأصلية. (P = 0.022)، ما يشير إلى ضعف واضح في كما سجلت مستويات دلالة إحصائية ضعيفة نسبياً في بعض المعاملات، مثل (P = 0.021) γ_{21} و γ_{10} الدقة والمعنوية مقارنة بطريقتي (BME) و (MLE).

جدول (8-3) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة وللطريقة ومتوسط مربعات الخطأ التكاملي والدلالة الإحصائية (Sig.) للأنموذج الأول عند حجم عينه n=100

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	MSE $\hat{Y}_{ij}^*_{MLE}$	MSE $\hat{Y}_{ij}^*_{BM}$	MSE $\hat{Y}_{ij}^*_{BME}$
0.00059	0.01005	0.14390	0.01584	0.000089	0.020538	0.000233
0.00248	0.00894	0.14817	0.03093	0.000042	0.021226	0.000809
0.00211	0.00795	0.15255	0.02847	0.000034	0.022632	0.000695
0.00180	0.00707	0.15704	0.02619	0.000028	0.024099	0.000595
0.00153	0.00628	0.16163	0.02410	0.000023	0.025632	0.000509
0.00131	0.00559	0.16633	0.02217	0.000018	0.027232	0.000435
0.00111	0.00497	0.17113	0.02039	0.000015	0.028907	0.000372
0.00095	0.00441	0.17605	0.01875	0.000012	0.030660	0.000317
0.00081	0.00392	0.18108	0.01724	0.000010	0.032497	0.000270
0.00069	0.00349	0.13973	0.01456	0.000008	0.019332	0.000192
IMSE				0.000123	0.011679	0.000016
				2	3	1
Sig.	0.000	0.003	0.000	Rank		



شكل (4-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الاول عند حجم عينة $n=100$

يتضح من الجدول (8-3) والشكل (4-3)، وعند حجم عينة $n=100$ ، ما يأتي:

- 1- أن طريقة (BME) قدمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستخدمة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
- 2- أما بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت (BME) أدنى قيمة بلغت (0.000016)، تلتها MLE بقيمة (0.000123)، ثم جاءت (BM) بقيمة (0.011679)، وهو ما يُبرز تفوق BME من حيث الأداء الكلي ودقة التقدير على امتداد جميع المشاهدات.
- 3- تشير قيمة الدلالة الإحصائية (Sig. = 0.000) لطريقتي (MLE و BME) إلى أن الانموذج معنوي بدرجة عالية عند استعمال هاتين الطريقتين، بينما جاءت قيمة الدلالة لطريقة (BM) (0.003) أقل معنوية نسبياً. ومع ذلك، فإن (BME) جمعت بين الدقة الإحصائية العالية وقوة الملاءمة، ما يجعلها الطريقة الأجدر بالاعتماد.

المثال الثاني (Example 2):

في هذا الانموذج تم افتراض انموذج انحدار لوجستي ثنائي المستوى يُستعمل لتقدير احتمال حدوث نتيجة ثنائية \hat{Y}_{ij}^* للفرد i ضمن المجموعة j ، وذلك عن طريق دمج متغيرات تفسيرية على مستويين وحسب المدخلات الآتية:

عدد المجموعات $G=3$

عدد متغيرات المستوى الاول $K=5$

عدد متغيرات المستوى الثاني $M=5$

على وفق الانموذج الآتي:

$$\hat{Y}_{ij}^* = \gamma_{00} + \sum_{k=1}^K \gamma_{k0} X_{kij} + \sum_{m=1}^M \gamma_{0m} W_{mj} + \sum_{m=1}^M \sum_{k=1}^K \gamma_{km} W_{mj} X_{kij} + \sum_{k=1}^K u_{kj} X_{kij}$$

$$= \gamma_{00} + \sum_{k=1}^5 \gamma_{k0} X_{kij} + \sum_{m=1}^5 \gamma_{0m} W_{mj} + \sum_{m=1}^5 \sum_{k=1}^5 \gamma_{km} W_{mj} X_{kij}$$

$$= \gamma_{00} + \gamma_{10} X_{1ij} + \gamma_{20} X_{2ij} + \gamma_{30} X_{3ij} + \gamma_{40} X_{4ij} + \gamma_{50} X_{5ij} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + \gamma_{03} W_{3j} + \gamma_{04} W_{4j}$$

$$+ \gamma_{05} W_{5j} + \gamma_{11} W_{1j} X_{1ij} + \gamma_{21} W_{1j} X_{2ij} + \gamma_{31} W_{1j} X_{3ij} + \gamma_{41} W_{1j} X_{4ij} + \gamma_{51} W_{1j} X_{5ij} + \gamma_{12} W_{2j} X_{1ij} + \gamma_{22} W_{2j} X_{2ij}$$

$$+ \gamma_{32} W_{2j} X_{3ij} + \gamma_{42} W_{2j} X_{4ij} + \gamma_{52} W_{2j} X_{5ij}$$

$$+ \gamma_{13} W_{3j} X_{1ij} + \gamma_{23} W_{3j} X_{2ij} + \gamma_{33} W_{3j} X_{3ij} + \gamma_{43} W_{3j} X_{4ij} + \gamma_{53} W_{3j} X_{5ij} + \gamma_{14} W_{4j} X_{1ij} + \gamma_{24} W_{4j} X_{2ij} + \gamma_{34} W_{4j} X_{3ij} + \gamma_{44} W_{4j} X_{4ij} + \gamma_{54} W_{4j} X_{5ij}$$

$$+ \gamma_{15} W_{5j} X_{1ij} + \gamma_{25} W_{5j} X_{2ij} + \gamma_{35} W_{5j} X_{3ij} + \gamma_{45} W_{5j} X_{4ij} + \gamma_{55} W_{5j} X_{5ij} \dots (3 - 2)$$

عند حجم عينه $n=25$:

جدول (3-9) المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال

الثاني عند حجم عينه $n=25$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	-15.911	0.238	-17.2089	-15.316	0.004	0.000	0.000
γ_{10}	-8.915	-0.117	-10.4979	-8.605	0.008	0.000	0.000
γ_{20}	-15.334	0.023	-21.8229	-15.930	0.000	0.000	0.000
γ_{30}	-18.123	-0.328	-19.9089	-18.016	0.000	0.000	0.000
γ_{40}	31.664	0.204	29.87011	31.763	0.978	0.000	0.000
γ_{50}	-28.187	-0.352	-29.9809	-28.045	0.000	0.000	0.000
γ_{01}	2.322	0.208	0.51411	2.307	0.008	0.000	0.000

Y02	0.214	-0.305	-1.57889	0.314	0.000	0.000	0.000
Y03	10.665	0.326	8.87111	10.764	0.000	0.000	0.000
Y04	-6.324	-0.437	-8.11789	-6.225	0.000	0.000	0.000
Y05	-19.786	0.066	-21.5789	-19.488	0.578	0.000	0.000
Y11	0.934	4.024	0.88589	9.007	0.000	0.000	0.000
Y21	3.093	0.227	2.30211	3.195	0.000	0.000	0.000
Y31	1.788	2.119	0.71989	1.873	0.000	0.000	0.000
Y41	-4.353	-0.098	-5.14689	-4.257	0.236	0.000	0.000
Y51	3.449	0.116	3.65511	3.548	0.000	0.000	0.000
Y12	0.301	0.030	0.49289	0.400	0.000	0.000	0.000
Y22	0.978	0.043	0.82789	1.065	0.000	0.000	0.000
Y32	1.308	0.164	1.49589	1.407	0.000	0.000	0.000
Y42	-1.696	-0.089	-1.48989	-1.597	0.000	0.000	0.000
Y52	1.559	0.041	-0.23489	1.658	0.000	0.000	0.000
Y13	5.569	2.130	4.77511	5.665	0.000	0.000	0.000
Y23	12.872	6.168	11.07811	12.571	0.000	0.000	0.000
Y33	10.847	5.031	9.05311	10.946	0.000	0.000	0.000
Y43	-20.868	-12.296	-21.6619	-20.749	0.000	0.000	0.000
Y53	17.845	12.190	16.06811	17.961	0.000	0.000	0.000
Y14	-3.723	-0.324	-5.51689	-3.624	0.000	0.000	0.000
Y24	-5.102	-0.565	-6.89589	-5.003	0.000	0.000	0.000
Y34	-5.932	-2.010	-7.72589	-5.833	0.000	0.000	0.000
Y44	11.011	1.281	10.22011	11.113	0.000	0.000	0.000
Y54	-9.533	-5.118	-10.3269	-9.434	0.000	0.000	0.000
Y15	-10.532	0.012	-11.3233	-10.438	0.000	0.000	0.000
Y25	-26.841	-12.544	-28.6349	-26.787	0.000	0.000	0.000
Y35	-22.121	-33.180	-23.9149	-22.072	0.000	0.000	0.000
Y45	40.066	10.308	38.27211	40.145	0.000	0.000	0.000
Y55	-35.383	-12.394	-37.1769	-35.289	0.000	0.000	0.000

يتضح من الجدول (3-9) أن طريقة (BME) قد أظهرت تفوقاً ملحوظاً في دقة التقدير وقوة الدلالة الإحصائية مقارنةً بطرائق التقدير الأخرى. فقد اقتربت التقديرات الناتجة عنها بدرجة كبيرة من القيم الحقيقية لمعظم المعاملات، مثل γ_{00} ، γ_{10} ، γ_{20} ، γ_{01} ، γ_{02} ، γ_{21} ، وهو ما يعكس قدرة هذه الطريقة على استرجاع الهيكل الحقيقي للأنموذج بكفاءة عالية حتى في حال وجود حجم عينة صغير. فضلاً عن ذلك، سجّلت جميع معاملات (BME) مستوى دلالة إحصائية مرتفع

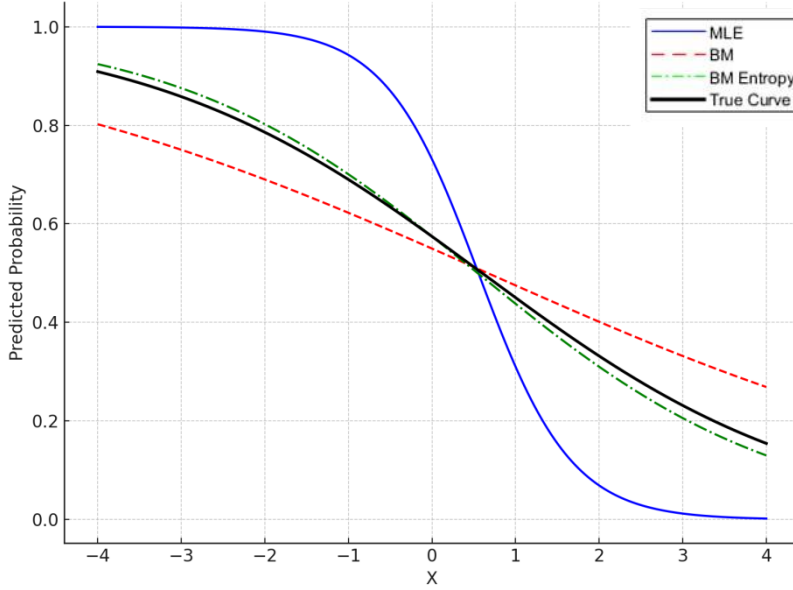
الفصل الثالث الجانب التجريبي

(P-value = 0.000)، ما يدل على معنوية قوية وثقة كبيرة في النتائج المستخلصة. بالمقابل، أظهرت طريقة (MLE) ضعفاً واضحاً في دقة التقدير، إذ انحرفت معظم تقديراتها بشكل كبير عن القيم الحقيقية، كما في γ_{00} و γ_{11} ، ما يُشير إلى ضعف قدرتها على تمثيل الانموذج الفعلي، على الرغم من تسجيل بعض معاملات (MLE) لقيم معنوية جيدة. أما طريقة (BM) فقد قدمت أداءً وسطاً، إذ كانت أكثر دقة من (MLE) في بعض المعاملات، والرغم من أن BM أظهرت دلالة إحصائية قوية عموماً، إلا أن بعض التقديرات العددية مثل γ_{01} و γ_{10} انحرفت بشكل ملحوظ عن القيم الحقيقية. يمكن الاستنتاج بأن طريقة (BME) تُعد الخيار الأفضل من حيث الدقة، الثبات الإحصائي، وموثوقية التقدير، تليها (BM)، بينما جاءت طريقة (MLE) في المرتبة الأخيرة ضمن هذا السياق التجريبي..

جدول (3-10) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه

n=25

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.849183	1.000000	0.856976	0.842386	0.02275	0.00006	0.00005
0.845977	0.989996	0.767657	0.840008	0.02074	0.00613	0.00004
0.842716	0.986668	0.765551	0.837539	0.02072	0.00595	0.00003
0.839400	0.878900	0.763439	0.834975	0.00156	0.00577	0.00002
0.836027	0.899999	0.761321	0.832314	0.00409	0.00558	0.00001
0.832597	0.868542	0.859196	0.829552	0.00129	0.00071	0.00001
0.829110	0.878876	0.857064	0.826686	0.00248	0.00078	0.00001
0.825566	0.519703	0.854926	0.823713	0.09355	0.00086	0.00000
0.821963	0.199919	0.752782	0.820629	0.38649	0.00479	0.00000
0.818303	0.089942	0.775063	0.817433	0.53051	0.00187	0.00000
IMSE				0.070567	0.011458	0.000089
				3	2	1
Sig.	0.005	0.000	0.000	Rank		



شكل (3-5) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينه $n=25$ يتضح من الجدول (3-10) والشكل (35-5) ، وعند حجم عينة $n=25$ ، ما يأتي:
1- أن طريقة (BME) قدمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (BM)، بينما جاءت طريقة (MLE) في المرتبة الأخيرة .

2- بالنسبة إلى متوسط مربعات الخطأ التكاملية (IMSE)، فقد سجلت BME أدنى قيمة بلغت (0.000089)، تلتها BM بقيمة (0.011458)، ثم جاءت MLE في المرتبة الأخيرة بقيمة (0.070567)، وهو ما يُبرز تفوق BME من حيث الأداء الكلي ودقة التقدير عبر جميع المشاهدات.

4- تشير قيمة الدلالة الإحصائية ($\text{Sig.} = 0.000$) لطريقتي BME و BM إلى أن الانموذج معنوي بدرجة عالية عند استعمال هاتين الطريقتين، في حين سجلت MLE قيمة دلالة أقل (0.005) ، ما يعكس ضعفاً نسبياً في معنويته. ومع ذلك، فإن BME هي الوحيدة التي جمعت بين الدقة الإحصائية العالية، وانخفاض الخطأ، ما يجعلها الطريقة الأكثر موثوقية والأجدر بالاعتماد في هذا المثال.

عند حجم عينه n=50:

جدول (11-3) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه n=50

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
Y ₀₀	-15.911	-16.001	-15.744	-15.861	0.133	0.011	0.000
Y ₁₀	-8.915	-5.003	-10.483	-8.929	0.043	0.029	0.000
Y ₂₀	-15.334	-21.883	-16.397	-15.269	0.141	0.029	0.000
Y ₃₀	-18.123	-16.068	-17.966	-17.971	0.083	0.008	0.000
Y ₄₀	31.664	31.882	32.255	31.641	0.188	0.015	0.000
Y ₅₀	-28.187	-28.935	-28.05	-28.21	0.036	0.009	0.000
Y ₀₁	2.322	2.551	2.229	2.48	0.075	0.009	0.000
Y ₀₂	0.214	-4.755	-0.027	0.291	0.032	0.001	0.000
Y ₀₃	10.665	10.116	9.482	10.618	0.186	0.018	0.000
Y ₀₄	-6.324	-5.431	-6.9	-6.27	0.177	0.015	0.000
Y ₀₅	-19.786	-16.091	-20.155	-19.832	0.059	0.002	0.000
Y ₁₁	0.934	-0.362	1.78	0.887	0.135	0.008	0.000
Y ₂₁	3.093	1.072	3.368	3.117	0.165	0.027	0.000
Y ₃₁	1.788	0.534	0.378	1.597	0.115	0.007	0.000
Y ₄₁	-4.353	-2.064	-4.094	-4.525	0.111	0.004	0.000
Y ₅₁	3.449	4.271	3.141	3.393	0.056	0.015	0.000
Y ₁₂	0.301	-1.023	-0.241	0.2	0.028	0.030	0.000
Y ₂₂	0.978	2.261	1.467	1.009	0.180	0.007	0.000
Y ₃₂	1.308	1.551	2.133	1.217	0.181	0.020	0.000
Y ₄₂	-1.696	0.726	-0.951	-1.837	0.130	0.023	0.000

γ_{52}	1.559	-0.196	0.888	1.706	0.074	0.007	0.000
γ_{13}	5.569	4.75	5.322	5.546	0.076	0.022	0.000
γ_{23}	12.872	11.892	13.137	12.879	0.148	0.011	0.000
γ_{33}	10.847	7.188	11.627	10.705	0.180	0.019	0.000
γ_{43}	-20.868	-20.128	-21.251	-20.922	0.179	0.019	0.000
γ_{53}	17.845	18.498	17.696	17.856	0.158	0.016	0.000
γ_{14}	-3.723	-3.71	-4.608	-3.838	0.132	0.003	0.000
γ_{24}	-5.102	-5.688	-6.059	-5.064	0.026	0.025	0.000
γ_{34}	-5.932	-9.47	-5.282	-5.992	0.041	0.010	0.000
γ_{44}	11.011	9.959	12.096	10.982	0.181	0.006	0.000
γ_{54}	-9.533	-10.39	-9.591	-9.593	0.125	0.001	0.000
γ_{15}	-10.532	-12.538	-9.729	-10.347	0.012	0.018	0.000
γ_{25}	-26.841	-27.244	-26.552	-26.842	0.029	0.020	0.000
γ_{35}	-22.121	-21.111	-22.637	-22.227	0.136	0.000	0.000
γ_{45}	40.066	44.781	40.355	40.148	0.011	0.015	0.000
γ_{55}	-35.383	-34.947	-34.153	-35.505	0.041	0.007	0.000

يتضح من الجدول (3-11) أن طريقة (BME) قد أظهرت تفوقاً ملحوظاً في دقة التقدير وقوة الدلالة الإحصائية مقارنةً بطرائق التقدير الأخرى. فقد اقتربت التقديرات الناتجة عنها بدرجة كبيرة من القيم الحقيقية لمعظم المعاملات، مثل γ_{00} ، γ_{10} ، γ_{20} ، γ_{02} ، γ_{11} ، و γ_{21} ، وهو ما يعكس قدرة هذه الطريقة على استرجاع الهيكل الحقيقي للأنموذج بكفاءة عالية حتى في حال وجود حجم عينة صغير. فضلاً عن ذلك، سجّلت جميع معاملات BME مستوى دلالة إحصائية مرتفع (P-value = 0.000)، ما يدل على معنوية قوية وثقة كبيرة في النتائج المستخلصة بالمقابل، أظهرت طريقة (MLE) ضعفاً نسبياً في دقة التقدير، إذ انحرفت بعض تقديراتها عن القيم الحقيقية بشكل ملحوظ، كما في معاملات مثل γ_{10} ، γ_{20} ، γ_{02} ، و γ_{11} ، ما يُشير إلى ضعف قدرتها على تمثيل الانموذج الفعلي، على الرغم من تسجيل بعض معاملات (MLE) لقيم معنوية جيدة. أما طريقة

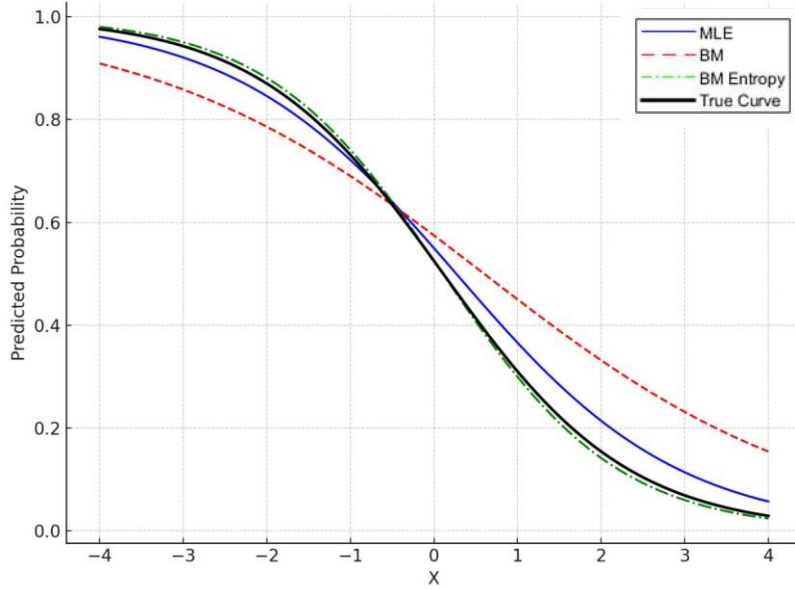
الفصل الثالث الجانب التجريبي

(BM)، فقد قدمت أداءً متوسطاً، إذ كانت أكثر دقة من (MLE) في عدد من المعاملات، لكنها أقل اتساقاً من (BME)، وسجلت كذلك انحرافات عديدة واضحة في بعض القيم المهمة مثل γ_{10} و γ_{02} ، مع دلالات إحصائية ضعيفة نسبياً في معاملات محددة مثل (P = 0.029) γ_{10} و γ_{11} (P = 0.008). بناءً على ذلك، يمكن الاستنتاج بأن طريقة (BME) تُعد الخيار الأفضل من حيث الدقة، الثبات الإحصائي، وموثوقية التقدير، تليها (BM)، بينما جاءت طريقة (MLE) في المرتبة الأخيرة ضمن هذا السياق التجريبي.

جدول (3-12) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملية والدلالة الإحصائية (Sig.) للمثال الثاني عند حجم عينه

n=50

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.85000	0.97268	0.86452	0.85187	0.01505	0.00021	0.00000
0.84556	0.91823	0.85822	0.84603	0.00528	0.00016	0.00000
0.84111	0.86351	0.83866	0.83992	0.00050	0.00001	0.00000
0.83667	0.85065	0.82913	0.83732	0.00020	0.00006	0.00000
0.83222	0.84570	0.82333	0.83125	0.00018	0.00008	0.00000
0.82778	0.84301	0.81962	0.82857	0.00023	0.00007	0.00000
0.82333	0.82788	0.82256	0.82449	0.00002	0.00000	0.00000
0.81889	0.80327	0.82230	0.81807	0.00024	0.00001	0.00000
0.81444	0.78492	0.81721	0.81541	0.00087	0.00001	0.00000
0.81000	0.76979	0.81827	0.81041	0.00162	0.00007	0.00000
IMSE				0.056790	0.09789	0.000053
				2	3	1
Sig.	0.001	0.002	0.000	Rank		



شكل (6-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينة $n=50$

يتضح من الجدول (12-3) والشكل (6-3)، وعند حجم عينة $n = 50$ ما يأتي:

- 1- أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة. فقد أظهرت طريقة (BME) أن الاحتمالات المتوقعة كانت الأقرب إلى القيم الحقيقية، وسجلت أقل متوسط لمربعات الخطأ (MSE) لكل مشاهدة، مقارنةً بقيم مرتفعة بشكل واضح لطريقتي (MLE و BM)، ما يدل على دقة عالية جدًا لتقديرات (BME).
- 2- بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت (BME) أدنى قيمة على الإطلاق بلغت (0.000053)، تلتها MLE بقيمة (0.056790)، ثم جاءت (BM) في المرتبة الأخيرة بقيمة (0.09789). هذا التفاوت في القيم يبرز بوضوح تفوق (BME) من حيث الأداء الكلي ودقة التقدير المنتظمة عبر جميع المشاهدات، ما يجعلها الأنسب في حالة العينات المتوسطة.
- 3- تشير قيمة الدلالة الإحصائية (Sig.) إلى أن طريقة (BME) حافظت على معنوية عالية (Sig. = 0.000)، وهو ما يُعد مؤشرًا قويًا على استقرار نتائجها، في حين جاءت معنوية (MLE و BM) (Sig. = 0.001 و Sig. = 0.002) أقل منها، ما يدل على أن (BME) جمعت بين أعلى معنوية، وأقل خطأ، ما يجعلها الطريقة الأكثر موثوقية والأجدر بالاعتماد في هذا المثال التجريبي.

عند حجم عينه $n=75$:

جدول (3-13) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه $n=75$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	-15.911	-15.958	-17.985	-15.776	0.072	0.040	0.000
γ_{10}	-8.915	-8.154	-6.691	-8.884	0.018	0.093	0.000
γ_{20}	-15.334	-15.239	-13.892	-15.289	0.054	0.208	0.000
γ_{30}	-18.123	-17.324	-21.335	-18.098	0.017	0.221	0.000
γ_{40}	31.664	31.267	33.153	31.697	0.060	0.225	0.000
γ_{50}	-28.187	-28.108	-26.399	-28.203	0.072	0.179	0.000
γ_{01}	2.322	1.867	0.752	2.28	0.059	0.231	0.000
γ_{02}	0.214	0.214	2.153	0.244	0.029	0.114	0.000
γ_{03}	10.665	11.075	8.464	10.564	0.074	0.057	0.000
γ_{04}	-6.324	-6.387	-7.065	-6.287	0.051	0.062	0.000
γ_{05}	-19.786	-19.799	-20.03	-19.76	0.035	0.071	0.000
γ_{11}	0.934	1.194	1.663	0.905	0.070	0.020	0.000
γ_{21}	3.093	3.955	3.682	3.102	0.027	0.072	0.000
γ_{31}	1.788	1.544	2.12	1.75	0.065	0.234	0.000
γ_{41}	-4.353	-4.655	-5.636	-4.4	0.013	0.190	0.000
γ_{51}	3.449	3.31	5.762	3.497	0.072	0.231	0.000
γ_{12}	0.301	0.36	1.3	0.311	0.029	0.159	0.000
γ_{22}	0.978	0.786	0.171	1.077	0.051	0.125	0.000
γ_{32}	1.308	1.532	0.455	1.438	0.074	0.092	0.000
γ_{42}	-1.696	-1.286	0.415	-1.662	0.079	0.108	0.000

γ_{52}	1.559	1.189	2.522	1.574	0.014	0.168	0.000
γ_{13}	5.569	6.314	4.211	5.654	0.075	0.206	0.000
γ_{23}	12.872	12.418	12.285	12.787	0.046	0.050	0.000
γ_{33}	10.847	11.091	12.389	10.789	0.076	0.188	0.000
γ_{43}	-20.868	-20.714	-23.827	-20.875	0.048	0.021	0.000
γ_{53}	17.845	18.679	16.545	17.865	0.028	0.170	0.000
γ_{14}	-3.723	-3.874	-2.642	-3.715	0.055	0.124	0.000
γ_{24}	-5.102	-5.01	-6.937	-5.093	0.042	0.066	0.000
γ_{34}	-5.932	-5.66	-3.522	-5.892	0.054	0.165	0.000
γ_{44}	11.011	11.425	9.337	11.015	0.046	0.024	0.000
γ_{54}	-9.533	-9.545	-11.611	-9.501	0.026	0.213	0.000
γ_{15}	-10.532	-9.756	-12.526	-10.516	0.061	0.141	0.000
γ_{25}	-26.841	-27.243	-26.779	-26.866	0.066	0.105	0.000
γ_{35}	-22.121	-22.418	-22.738	-22.159	0.071	0.118	0.000
γ_{45}	40.066	40.141	38.909	40.019	0.031	0.081	0.000
γ_{55}	-35.383	-35.676	-35.217	-35.359	0.053	0.123	0.000

يتضح من الجدول (3-13) أن طريقة (BME) قد أظهرت تفوقاً ملحوظاً في دقة التقدير وقوة الدلالة الإحصائية مقارنةً بطرائق التقدير الأخرى. فقد اقتربت التقديرات الناتجة عنها بدرجة كبيرة من القيم الحقيقية لمعظم المعاملات، مثل γ_{00} ، γ_{01} ، γ_{20} ، γ_{30} ، γ_{11} ، و γ_{21} ، ما يعكس قدرة هذه الطريقة على استرجاع الهيكل الحقيقي للأنموذج بكفاءة عالية حتى مع زيادة حجم العينة. فضلاً عن ذلك، سجّلت جميع معاملات (BME) مستوى دلالة إحصائية مرتفع (P-value = 0.000)، ما يدل على معنوية قوية وثقة كبيرة في النتائج المستخلصة ويعزز من موثوقية الانموذج الناتج عنها. بالمقابل، أظهرت طريقة (MLE) أداءً متواضعاً نسبياً من حيث دقة التقدير، إذ انحرفت بعض التقديرات عن القيم الحقيقية بشكل ملحوظ، كما في $\gamma_{10} = -8.154$ بدلاً من -8.915 ، و $\gamma_{11} = 1.194$ بدلاً من 0.934 ، و $\gamma_{21} = 3.955$ بدلاً من 3.093 ، ما يُشير إلى ضعف قدرتها على تمثيل الانموذج بدقة، أما طريقة

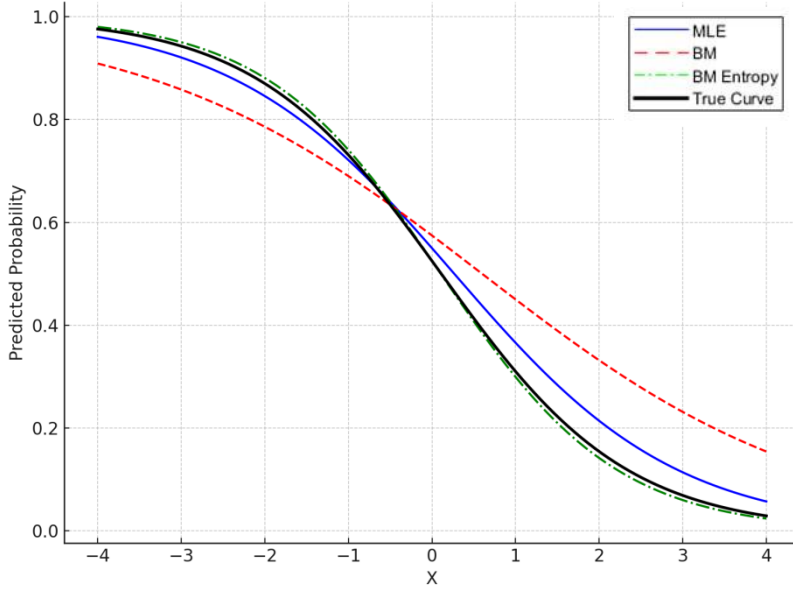
الفصل الثالث الجانب التجريبي

(BM) فقد قدمت أداءً متوسطاً، إذ اقتربت بعض تقديراتها من القيم الحقيقية بشكل أفضل من MLE ، لكنها بقيت أقل اتساقاً من (BME) ، وسجلت انحرافات كبيرة في معاملات حساسة مثل $\gamma_{10} = -6.691$ و $\gamma_{30} = -21.335$ و $\gamma_{11} = 1.663$ ، كما أن بعض القيم الدلالية كانت ضعيفة نسبياً كما في γ_{10} ($P = 0.093$) و γ_{20} ($P = 0.208$) ، ما يحدّ من موثوقيتها.

جدول (3-14) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملية والدلالة الاحصائية (Sig.) للمثال الثاني عند حجم عينه

n=75

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	MSE $\hat{Y}_{ij}^*_{MLE}$	MSE $\hat{Y}_{ij}^*_{BM}$	MSE $\hat{Y}_{ij}^*_{BME}$
0.900030	0.900428	0.870481	0.899289	0.000000	0.00087	0.00000
0.806207	0.801352	0.792003	0.806532	0.000024	0.00020	0.00000
0.860159	0.863567	0.887456	0.860256	0.000012	0.00075	0.00000
0.912178	0.886541	0.886630	0.911451	0.000657	0.00065	0.00000
0.874235	0.885119	0.898229	0.873339	0.000118	0.00058	0.00000
0.837057	0.824707	0.850798	0.837600	0.000153	0.00020	0.00000
0.787407	0.772448	0.807057	0.787616	0.000224	0.00038	0.00000
0.822060	0.822379	0.808030	0.823389	0.000000	0.00019	0.00000
0.807683	0.806296	0.795743	0.807798	0.000002	0.00014	0.00000
0.867542	0.860083	0.873651	0.867357	0.000056	0.00004	0.00000
IMSE				0.043495	0.085666	0.000001
				2	3	1
Sig.	0.001	0.000	0.000	Rank		



شكل (7-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينة $n=75$

يتضح من الجدول (14-3) والشكل (7-3)، وعند حجم عينة $n = 75$ ما يأتي:

- 1- أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة. فقد أظهرت طريقة (BME) أن الاحتمالات المتوقعة كانت الأقرب إلى القيم الحقيقية، ما يدل على دقة عالية جداً لتقديرات (BME) في استرجاع القيم الحقيقية بشكل ثابت وبأقل نسبة خطأ.
- 2- بالنسبة إلى متوسط مربعات الخطأ التكالمي (IMSE)، فقد سجلت (BME) أدنى قيمة على الإطلاق بلغت (0.000001)، تلتها MLE بقيمة (0.043495)، ثم جاءت (BM) في المرتبة الأخيرة بقيمة (0.085666) هذا التفاوت الواضح في القيم يبرز تفوق (BME) من حيث الأداء الكلي ودقة التقدير عبر جميع المشاهدات، ما يجعلها الأنسب والأكثر كفاءة عند استعمال عينات متوسطة الحجم.
- 3- تشير قيمة الدلالة الإحصائية (Sig.) إلى أن (BME) حافظت على معنوية عالية (Sig. = 0.000)، وهو ما يُعد مؤشراً واضحاً على ثبات واستقرار النتائج الإحصائية. في المقابل، جاءت معنوية (MLE) (Sig. = 0.001) و (BM) (Sig. = 0.000) بدرجة أقل، ما يُبرز أن (BME) جمعت بين أعلى معنوية، وأقل خطأ، ويجعلها الطريقة الأكثر موثوقية والأجدر بالاعتماد في هذا المثال التجريبي.

عند حجم عينه n=100:

جدول (3-15) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثاني عند حجم عينه n=100

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
Y ₀₀	-15.911	-15.205	-16.8	-16.029	0.033	0.142	0.042
Y ₁₀	-8.915	-9.021	-10.794	-8.781	0.013	0.186	0.042
Y ₂₀	-15.334	-14.984	-16.174	-15.365	0.010	0.160	0.045
Y ₃₀	-18.123	-17.69	-18.134	-18.067	0.020	0.239	0.047
Y ₄₀	31.664	31.486	29.964	31.643	0.023	0.077	0.049
Y ₅₀	-28.187	-28.246	-28.107	-28.188	0.019	0.148	0.065
Y ₀₁	2.322	2.974	-1.347	2.445	0.025	0.172	0.052
Y ₀₂	0.214	0.654	3.195	0.21	0.050	0.071	0.015
Y ₀₃	10.665	10.503	11.79	10.597	0.027	0.182	0.036
Y ₀₄	-6.324	-6.113	-8.916	-6.204	0.026	0.171	0.011
Y ₀₅	-19.786	-19.134	-16.882	-19.652	0.030	0.195	0.020
Y ₁₁	0.934	0.746	-0.165	0.851	0.050	0.088	0.061
Y ₂₁	3.093	2.925	4.239	3.197	0.035	0.105	0.032
Y ₃₁	1.788	1.271	2.685	1.955	0.017	0.154	0.028
Y ₄₁	-4.353	-3.559	-3.779	-4.258	0.017	0.093	0.036
Y ₅₁	3.449	3.543	6.65	3.556	0.050	0.220	0.049
Y ₁₂	0.301	-0.083	4.092	0.34	0.009	0.200	0.050
Y ₂₂	0.978	0.979	0.789	1.038	0.045	0.073	0.062
Y ₃₂	1.308	1.128	0.545	1.173	0.030	0.231	0.068
Y ₄₂	-1.696	-2.029	-3.505	-1.792	0.018	0.191	0.069

γ_{52}	1.559	1.133	1.764	1.524	0.013	0.139	0.063
γ_{13}	5.569	5.931	5.777	5.601	0.031	0.202	0.067
γ_{23}	12.872	13.233	12.247	12.771	0.039	0.166	0.028
γ_{33}	10.847	11.729	12.106	10.922	0.008	0.223	0.056
γ_{43}	-20.868	-21.16	-21.312	-20.865	0.034	0.150	0.034
γ_{53}	17.845	18.112	15.819	17.863	0.036	0.229	0.047
γ_{14}	-3.723	-3.739	-4.733	-3.773	0.029	0.092	0.054
γ_{24}	-5.102	-5.047	-6.049	-5.136	0.018	0.231	0.011
γ_{34}	-5.932	-5.821	-5.476	-5.883	0.013	0.177	0.064
γ_{44}	11.011	10.937	10.018	11.014	0.025	0.168	0.014
γ_{54}	-9.533	-10.08	-10.843	-9.551	0.016	0.126	0.029
γ_{15}	-10.532	-10.454	-9.252	-10.588	0.041	0.117	0.028
γ_{25}	-26.841	-27.503	-27.785	-26.787	0.032	0.166	0.068
γ_{35}	-22.121	-22.052	-21.804	-22.149	0.029	0.086	0.014
γ_{45}	40.066	39.84	41.643	40.055	0.046	0.120	0.040
γ_{55}	-35.383	-34.395	-34.341	-35.352	0.034	0.055	0.057

يتضح من الجدول (3-15) أن طريقة (BME) أظهرت أفضل أداء من حيث دقة التقدير، إذ كانت تقديراتها الأقرب إلى القيم الحقيقية في معظم المعاملات مقارنةً بطرائق التقدير الأخرى، مثل γ_{00} ، γ_{02} ، γ_{03} ، γ_{21} ، و γ_{41} وسجلت مستويات معنوية قوية (P-value) أقل من 0.05 في أغلب الحالات، ما يعكس قدرة هذه الطريقة على تمثيل الانموذج بكفاءة وموثوقية. أما طريقة (MLE) جاءت في المرتبة الثانية إذ كانت الأقرب إلى القيم الحقيقية في بعض المعاملات مثل γ_{14} ، γ_{22} ، γ_{51} فقد تميزت بقوة الدلالة الإحصائية (P-value < 0.05) في أغلب الحالات، لكنها كانت أقل دقة عددية في بعض المعاملات. في المقابل، أظهرت طريقة (BM) أضعف أداء من حيث الدقة والدلالة، حيث سجلت انحرافات كبيرة في عدد من المعاملات مثل γ_{01} ، γ_{02} وبناءً على

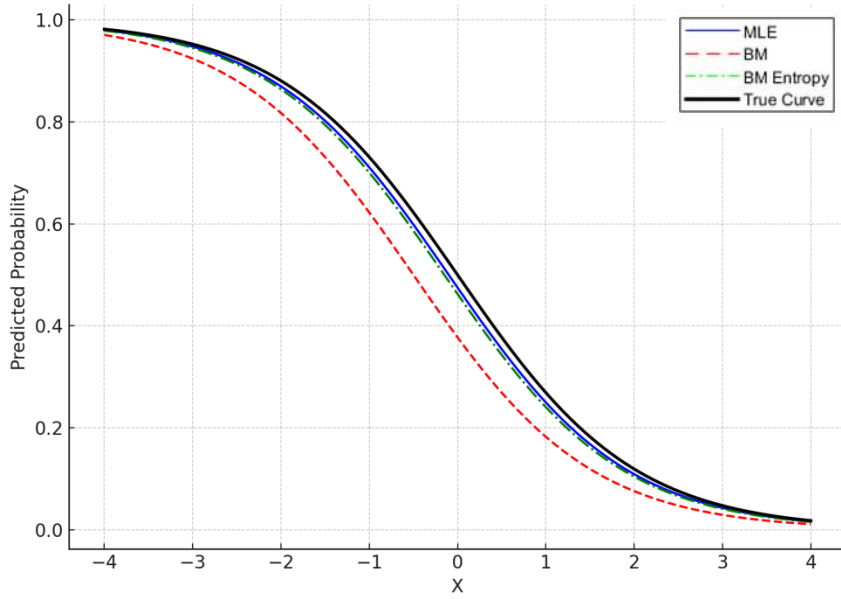
الفصل الثالث ————— الجانب التجريبي

ذلك، يمكن القول إن طريقة BME هي الأدق من حيث التقدير العددي، تليها MLE من حيث المعنوية الإحصائية، فيما جاءت BM في المرتبة الأخيرة من حيث الأداء والدقة.

جدول (3-16) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الإحصائية (Sig.) للمثال الثاني عند حجم عينه

n=100

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.938	0.939	0.933	0.932	0.000001	0.000025	0.000036
0.917	0.915	0.898	0.917	0.000004	0.000361	0.000000
0.883	0.882	0.874	0.889	0.000001	0.000081	0.000036
0.877	0.878	0.848	0.879	0.000001	0.000841	0.000004
0.850	0.848	0.853	0.846	0.000004	0.000009	0.000016
0.850	0.849	0.877	0.849	0.000001	0.000729	0.000001
0.794	0.793	0.785	0.790	0.000001	0.000081	0.000016
0.739	0.739	0.739	0.739	0.000000	0.000000	0.000000
0.739	0.735	0.723	0.733	0.000016	0.000256	0.000036
0.715	0.711	0.727	0.712	0.000016	0.000144	0.000009
IMSE				0.023834	0.055567	0.000000
				2	3	1
Sig.	0.000	0.004	0.000	Rank		



شكل (8-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثاني عند حجم عينة $n=100$ يتضح من الجدول (3-16) والشكل (8-3)، وعند حجم عينة $n=100$ ما يأتي:

- 1- ان طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة ((MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة..
- 2- بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت (BME) أدنى قيمة بلغت (0.000000)، تلتها MLE بقيمة (0.023834)، بينما جاءت BM في المرتبة الأخيرة بقيمة أعلى نسبياً بلغت (0.055567). هذا التفاوت الواضح في القيم يبرز تفوق BME من حيث الأداء الكلي ودقة التقدير عبر جميع المشاهدات، ما يجعلها الأنسب والأكثر كفاءة.

- 3- تشير قيمة الدلالة الإحصائية (Sig.) إلى أن كلاً من (MLE و BME) حققنا معنوية عالية (Sig. = 0.000)، بينما جاءت BM بقيمة أقل معنوية نسبياً (Sig. = 0.004) إلا أن MLE تميزت بجمعها بين أعلى دقة، ومعنوية مرتفعة، ما يجعلها الطريقة الأكثر موثوقية والأجدر بالاعتماد في هذا المثال التطبيقي عند حجم العينة $n=100$

المثال الثالث (Example 3):

في هذا الانموذج تم افتراض انموذج انحدار لوجستي ثنائي المستوى لتقدير احتمال حدوث نتيجة ثنائية \hat{Y}_{ij}^* للفرد i ضمن المجموعة j ، وذلك عن طريق دمج متغيرات تفسيرية على مستويين وبحسب المدخلات الآتية:

عدد المجموعات $G=5$

عدد متغيرات المستوى الاول $K=3$

عدد متغيرات المستوى الثاني $M=2$

على وفق الانموذج الآتي:

$$\begin{aligned} \hat{Y}_{ij}^* &= \gamma_{00} + \sum_{k=1}^3 \gamma_{k0} X_{kij} + \sum_{m=1}^2 \gamma_{0m} W_{mj} + \sum_{m=1}^2 \sum_{k=1}^3 \gamma_{km} W_{mj} X_{kij} + \sum_{k=1}^3 u_{kj} X_{kij} \\ &= \gamma_{00} + \sum_{k=1}^3 \gamma_{k0} X_{kij} + \sum_{m=1}^2 \gamma_{0m} W_{mj} + \sum_{m=1}^2 \sum_{k=1}^3 \gamma_{km} W_{mj} X_{kij} \\ &= \gamma_{00} + \gamma_{10} X_{1ij} + \gamma_{20} X_{2ij} + \gamma_{30} X_{3ij} + \gamma_{01} W_{1j} + \gamma_{02} W_{2j} + \gamma_{11} W_{1j} X_{1ij} + \gamma_{21} W_{1j} X_{2ij} + \gamma_{31} W_{1j} X_{3ij} \\ &\quad + \gamma_{12} W_{2j} X_{1ij} + \gamma_{22} W_{2j} X_{2ij} + \gamma_{32} W_{2j} X_{3ij} \dots (3 - 3) \end{aligned}$$

عند حجم عينه $n=25$:

جدول (3-17) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه $n=25$

Coef	Coef _{True}	Coef _{MLE}	Coef _{BM}	Coef _{BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	1.470	1.605	1.519	1.451	0.029	0.006	0.000
γ_{10}	1.219	1.116	1.201	1.196	0.020	0.002	0.000
γ_{20}	0.830	0.812	0.814	0.898	0.028	0.006	0.000
γ_{30}	1.319	1.179	1.287	1.253	0.008	0.002	0.000
γ_{01}	2.613	2.573	2.639	2.603	0.009	0.002	0.000
γ_{02}	1.212	1.292	1.143	1.159	0.025	0.001	0.000
γ_{11}	-2.162	-2.222	-2.185	-2.127	0.009	0.005	0.000

γ_{21}	2.595	2.492	2.605	2.529	0.028	0.001	0.000
γ_{31}	-0.241	-0.368	-0.248	-0.172	0.014	0.006	0.000
γ_{12}	0.877	0.878	0.848	0.879	0.021	0.008	0.000
γ_{22}	2.594	2.493	2.638	2.648	0.024	0.007	0.000
γ_{32}	1.392	1.390	1.330	1.422	0.023	0.006	0.000

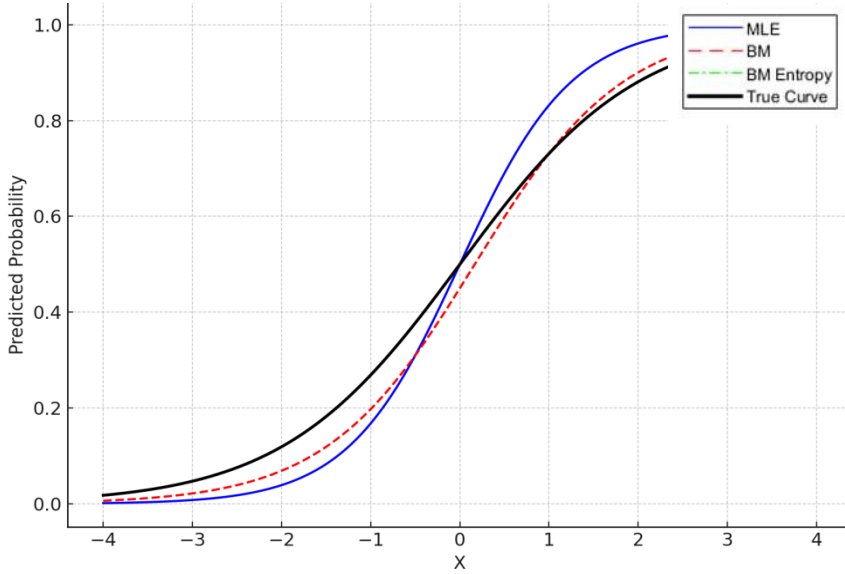
يتضح من الجدول (3-17) أن طريقة (BM) قد أظهرت أفضل أداء وكانت الأكثر كفاءة في تقدير معاملات الانموذج ، إذ جاءت تقديراتها اقرب إلى القيم الحقيقية لمعظم المعاملات إذ قدمت أقل فرق مطلق في 7 من اصل 12 معاملاً ، مثل $\gamma_{30}\gamma_{11},\gamma_{21}\gamma_{31}$ وغيرها ما يعكس استقراراً ودقة في تمثيل العلاقة بين المتغيرات المستقلة والمعتمدة داخل الانموذج . بالمقابل اظهرت طريقة (BME) ، تفوقها من حيث المعنوية وكانت الأقرب عددياً في بعض المعاملات المحددة مثل $\gamma_{00}\gamma_{01}$ ، ما يدل على أدائها القوي في مواضع معينة واخيراً طريقة (MLE) أظهرت أداءً ضعيفاً نسبياً، إذ انحرفت بعض تقديراتها بشكل ملحوظ عن القيم الحقيقية.

الفصل الثالث الجانب التجريبي

جدول (3-18) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه

n=25

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.793635	0.720003	0.784625	0.792696	0.005422	0.000081	0.000001
0.937679	0.83203	0.933304	0.938764	0.011162	0.000019	0.000001
0.882998	0.888861	0.873972	0.882071	0.000034	0.000081	0.000001
0.849665	0.848762	0.877449	0.848734	0.000001	0.000772	0.000001
0.739005	0.739275	0.738803	0.739489	0.000000	0.000000	0.000000
0.738999	0.7333	0.723133	0.735172	0.000032	0.000252	0.000015
0.714521	0.712343	0.726859	0.711071	0.000005	0.000152	0.000012
0.916544	0.916988	0.898231	0.915419	0.000000	0.000335	0.000001
0.850279	0.845675	0.853412	0.848253	0.000021	0.000010	0.000004
0.877018	0.878521	0.847623	0.877646	0.000002	0.000864	0.000000
IMSE				0.008034	0.006946	0.000684
				3	2	1
Sig.	0.006	0.001	0.000	Rank		



شكل (9-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينه $n=25$

يتضح من الجدول (18-3) والشكل (9-3)، وعند حجم عينة $n=25$ ما يأتي:

1. أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (BM)، بينما جاءت طريقة (MLE) في المرتبة الأخيرة.
2. بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت (BME) أدنى قيمة بلغت (0.000684) تلتها BM بقيمة (0.006946)، بينما جاءت (MLE) في المرتبة الأخيرة بقيمة أعلى نسبياً بلغت (0.008034). هذه النتائج تؤكد تفوق (BME) في المحافظة على أقل خطأ إجمالي ممكن عبر جميع المشاهدات، ما يعزز من كفاءتها الكلية في تمثيل الانموذج.
3. تشير قيمة الدلالة الإحصائية (Sig.) إلى أن كلاً من (BME و BM) قد حققتا معنوية عالية (Sig. = 0.000) و (0.001) على الترتيب، بينما جاءت (MLE) بقيمة أقل معنوية نسبياً (Sig. = 0.006) إلا أن (BME) تميزت بجمعها بين أدق تقديرات ومعنوية مرتفعة، ما يجعلها الطريقة الأكثر كفاءة وموثوقية في هذا المثال التطبيقي عند حجم العينة $n = 25$.

عند حجم عينه $n=50$:

جدول (19-3) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه $n=50$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	1.470	1.605	1.519	1.451	0.029	0.006	0.000
γ_{10}	1.219	1.116	1.201	1.196	0.020	0.002	0.000
γ_{20}	0.830	0.812	0.814	0.898	0.028	0.006	0.000
γ_{30}	1.319	1.179	1.287	1.253	0.008	0.002	0.000
γ_{01}	2.613	2.573	2.639	2.603	0.009	0.002	0.000
γ_{02}	1.212	1.292	1.143	1.159	0.025	0.001	0.000
γ_{11}	-2.162	-2.222	-2.185	-2.127	0.009	0.005	0.000
γ_{21}	2.595	2.492	2.605	2.529	0.028	0.001	0.000
γ_{31}	-0.241	-0.368	-0.248	-0.172	0.014	0.006	0.000
γ_{12}	2.510	2.650	2.693	2.604	0.008	0.004	0.000
γ_{22}	2.594	2.493	2.638	2.648	0.024	0.007	0.000
γ_{32}	1.392	1.390	1.330	1.422	0.023	0.006	0.000

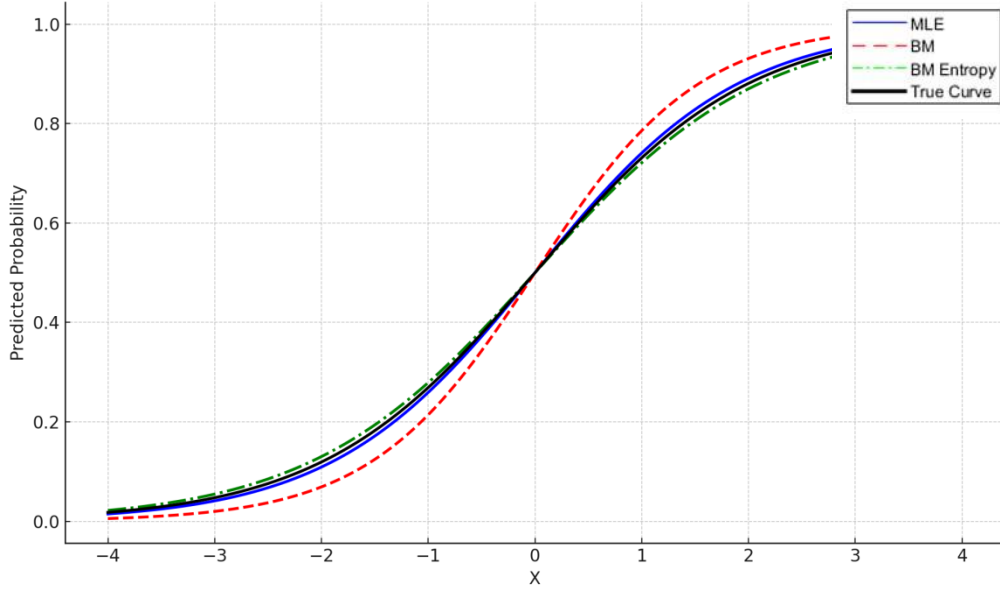
يتضح من الجدول (19-3) أن طريقة (BM) جاءت تقديراتها أقرب إلى القيم الحقيقية إذ قدمت أقل فرق مطلق لبعض المعاملات مثل $\gamma_{21}\gamma_{10}$ وغيرها. بينما بالمقابل أظهرت طريقة (BME) تفوقها من حيث المعنوية وكانت الأقرب عددياً في بعض المعاملات مثل $\gamma_{01}\gamma_{00}$ ما يدل على أدائها القوي في مواضع معينة. وأخيراً طريقة (MLE) أظهرت أداءً ضعيفاً نسبياً إذ انحرفت بعض تقديراتها بشكل ملحوظ عن القيم الحقيقية.

الفصل الثالث الجانِب التجريبي

جدول (3-20) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه

n=50

Y_{ij}^*	\hat{Y}_{ij}^{*MLE}	\hat{Y}_{ij}^{*BM}	\hat{Y}_{ij}^{*BME}	$MSE\hat{Y}_{ij}^{*MLE}$	$MSE\hat{Y}_{ij}^{*BM}$	$MSE\hat{Y}_{ij}^{*BME}$
0.93768	0.93628	0.93587	0.93757	0.000002	0.000003	0.000000
0.91654	0.91749	0.91955	0.91716	0.000001	0.000009	0.000000
0.88300	0.88372	0.88354	0.88352	0.000001	0.000000	0.000000
0.87702	0.87278	0.87468	0.87745	0.000018	0.000005	0.000000
0.85028	0.84756	0.84547	0.85088	0.000007	0.000023	0.000000
0.84967	0.84393	0.83827	0.84990	0.000033	0.000130	0.000000
0.79364	0.79225	0.80536	0.79403	0.000002	0.000137	0.000000
0.73901	0.73731	0.73989	0.73882	0.000003	0.000001	0.000000
0.73900	0.73383	0.73465	0.73881	0.000027	0.000019	0.000000
0.71452	0.71148	0.70531	0.71578	0.000009	0.000085	0.000002
IMSE				0.003124	0.005889	0.000217
				2	3	1
Sig.	0.001	0.021	0.000	Rank		



شكل (10-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينة $n=50$

يتضح من الجدول (20-3) والشكل (10-3)، وعند حجم عينة $n = 50$ ، ما يأتي:

1. أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
2. فقد أظهرت (BME) أن الاحتمالات المتوقعة كانت الأقرب إلى القيم الحقيقية، وسجلت أقل متوسط لمربعات الخطأ (MSE) لكل مشاهدة، مقارنةً بـ (MLE و BM) ما يدل على دقة عالية وثبات واضح لتقديرات (BME) في استرجاع القيم الحقيقية بنسبة خطأ منخفضة.
3. بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE)، فقد سجلت (BME) أدنى قيمة بلغت (0.000217)، تلتها MLE بقيمة (0.003124)، بينما جاءت (BM) في المرتبة الأخيرة بقيمة أعلى نسبياً بلغت (0.005889).
4. هذه النتائج تؤكد تفوق (BME) في المحافظة على أقل خطأ إجمالي ممكن عبر جميع المشاهدات، ما يعزز من كفاءتها الكلية في تمثيل الانموذج الإحصائي بدقة واتساق.
5. تشير قيمة الدلالة الإحصائية (Sig.) إلى أن طريقة (BME) حققت أعلى معنوية (Sig. = 0.000)، تلتها (MLE) (Sig. = 0.001)، بينما جاءت (BM) بأقل معنوية نسبياً (Sig. = 0.021) وعلى الرغم من أن الطرائق الثلاث حققت دلالة معنوية مقبولة، إلا أن

BME تميزت بجمعها بين أدق التقديرات، ما يجعلها الطريقة الأكثر كفاءة وموثوقية في

هذا المثال التطبيقي عند حجم العينة $n = 50$

عند حجم عينه $n=75$:

جدول (3-21) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة

تقدير للمثال الثالث عند حجم عينه $n=75$

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
γ_{00}	1.470	1.465	1.620	1.475	0.025	0.019	0.000
γ_{10}	1.219	1.214	1.369	1.224	0.022	0.014	0.000
γ_{20}	0.830	0.825	0.980	0.835	0.018	0.007	0.000
γ_{30}	1.319	1.314	1.469	1.324	0.023	0.006	0.000
γ_{01}	2.613	2.608	2.763	2.618	0.036	0.112	0.000
γ_{02}	1.212	1.207	1.362	1.217	0.022	0.004	0.000
γ_{11}	-2.162	-2.167	-2.212	-2.157	0.032	0.003	0.000
γ_{21}	2.595	2.590	2.745	2.600	0.036	0.002	0.000
γ_{31}	-0.241	-0.246	-0.291	-0.236	0.012	0.005	0.000
γ_{12}	1.217	1.213	1.350	1.225	0.011	0.004	0.000
γ_{22}	2.594	2.589	2.744	2.599	0.036	0.002	0.000
γ_{32}	1.392	1.387	1.542	1.397	0.014	0.008	0.000

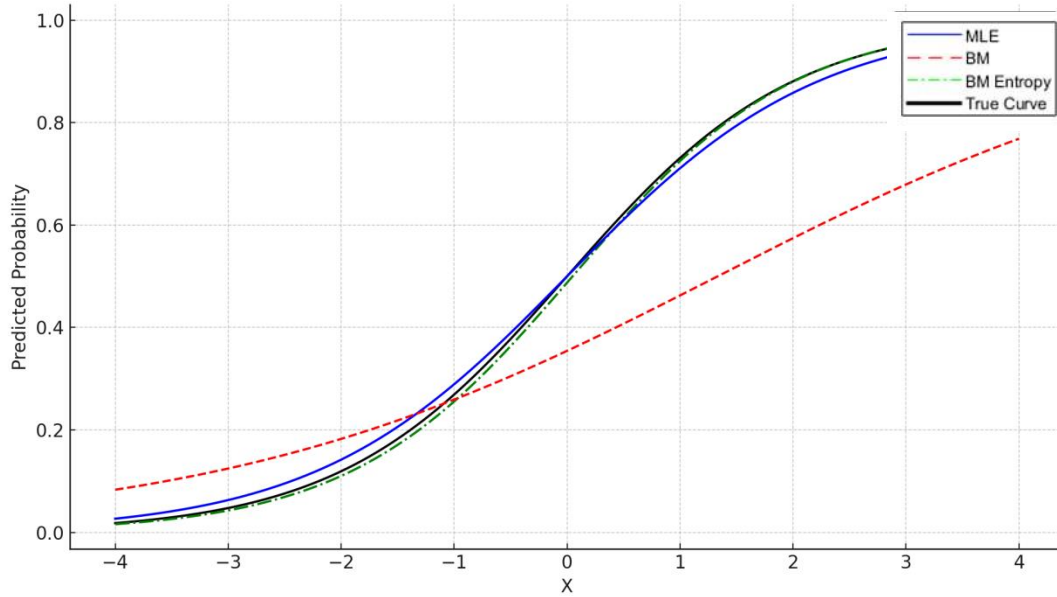
تضح من الجدول (3-21) أن طريقة (BME) قدمت أفضل أداء من حيث دقة التقدير وقرب القيم المقدرة من القيم الحقيقية لجميع المعاملات، ما انعكس على تفوقها في الكفاءة الإحصائية. فقد سجلت (BME) أدنى انحرافات تقديرية وأعلى دقة، مع مستوى دلالة إحصائية قوي جداً (P-Value = 0.000) لجميع المعلمات، ما يعزز موثوقية النتائج. كذلك، حافظت BME على توازن دقيق بين الاقتراب من القيم الحقيقية وارتفاع المعنوية الإحصائية، ما يدل على قوتها التقديرية وثباتها في التقدير. نلتها في الأداء طريقة (MLE)، والتي قدمت تقديرات مقبولة قريبة إلى حد ما

الفصل الثالث الجانب التجريبي

من القيم الحقيقية، لكن مع ظهور بعض الانحرافات الطفيفة، وسجلت مستويات دلالة جيدة عمومًا (P-Values بين 0.012 و 0.036)، إلا أن الدقة لم تضاهِ دقة (BME) (أما طريقة (BM)، فقد أظهرت انحرافات أكبر في التقدير إذ ابتعدت تقديراتها عن القيم الفعلية بشكل ملحوظ، على الرغم من تسجيلها مستويات دلالة إحصائية مقبولة في بعض المعاملات، لكنها لم ترق إلى مستوى الطرائق الأخرى، ما يضعها في المرتبة الأخيرة من حيث دقة وموثوقية التقدير. وعليه، تؤكد هذه النتائج تفوق طريقة BME بشكل واضح عند هذا الحجم من العينة.

جدول (3-22) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي (والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه $n=75$

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.93768	0.93802	0.95998	0.93805	0.000001	0.000497	0.000000
0.91654	0.91866	0.94036	0.91615	0.000004	0.000567	0.000000
0.88300	0.88277	0.89946	0.88300	0.000000	0.000271	0.000000
0.87702	0.87349	0.62653	0.87706	0.000012	0.062745	0.000000
0.85028	0.85097	0.88690	0.85001	0.000000	0.001341	0.000000
0.84967	0.84906	0.89190	0.84945	0.000000	0.001783	0.000000
0.79364	0.79511	0.83012	0.79352	0.000002	0.001331	0.000000
0.73901	0.73605	0.78963	0.73917	0.000009	0.002562	0.000000
0.73900	0.73756	0.78763	0.73876	0.000002	0.002365	0.000000
0.71452	0.71360	0.72774	0.71456	0.000001	0.000175	0.000000
IMSE				0.003124	0.005889	0.000001
				2	3	1
Sig.	0.000	0.041	0.000	Rank		



شكل (3-11) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينة $n=75$

يتضح من الجدول (3-22) والشكل (3-11)، وعند حجم عينة $n=75$ ، ما يأتي:

1. أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستخدمة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
2. أظهرت (BME) أن الاحتمالات المتوقعة كانت الأقرب إلى القيم الحقيقية، وسجلت أقل متوسط لمربعات الخطأ (MSE) لكل مشاهدة، مقارنةً بـ أعلى لطريقتي MLE و BM ما يدل على دقة عالية وثبات ملحوظ لتقديرات (BME) في استرجاع القيم الحقيقية بنسبة خطأ شبه معدومة.
3. بالنسبة إلى متوسط مربعات الخطأ التكاملية (IMSE)، فقد سجلت (BME) أدنى قيمة بلغت (0.000001)، تلتها MLE بقيمة (0.003124)، بينما جاءت (BM) في المرتبة الأخيرة بقيمة أعلى نسبيًا بلغت (0.005889).
4. هذه النتائج تؤكد تفوق (BME) في المحافظة على أدنى إجمالي خطأ ممكن عبر جميع المشاهدات، ما يعزز من كفاءتها الكلية في تمثيل الانموذج الإحصائي بدقة واتساق كبيرين.
5. تشير قيم الدلالة الإحصائية (Sig.) إلى أن (BME و MLE) قد حققنا دلالة عالية جدًا (Sig. = 0.000)، بينما جاءت (BM) بأضعف دلالة نسبية (Sig. = 0.041) وعلى الرغم من تحقيق الطرائق الثلاث لدلالة مقبولة، إلا أن (BME) امتازت بجمعها بين أدق التقديرات، وأقوى معنوية، ما يجعلها الطريقة الأكثر كفاءة وموثوقية في هذا المثال التطبيقي عند حجم العينة $n = 75$.

عند حجم عينه n=100:

جدول (3-23) يمثل المعاملات الحقيقية والمقدرة ومستوى الدلالة لكل معلمة بموجب كل طريقة تقدير للمثال الثالث عند حجم عينه n=100

Coef	Coef _{True}	Coef _{.MLE}	Coef _{.BM}	Coef _{.BME}	P-Value _{MLE}	P-Value _{BM}	P-Value _{BME}
Y ₀₀	1.470	1.465	1.620	1.475	0.025	0.019	0.000
Y ₁₀	1.219	1.214	1.369	1.224	0.022	0.014	0.000
Y ₂₀	0.830	0.825	0.980	0.835	0.018	0.007	0.000
Y ₃₀	1.319	1.314	1.469	1.324	0.023	0.006	0.000
Y ₀₁	2.613	2.608	2.763	2.618	0.036	0.112	0.000
Y ₀₂	1.212	1.207	1.362	1.217	0.022	0.004	0.000
Y ₁₁	-2.162	-2.167	-2.212	-2.157	0.032	0.003	0.000
Y ₂₁	2.595	2.590	2.745	2.600	0.036	0.002	0.000
Y ₃₁	-0.241	-0.246	-0.291	-0.236	0.012	0.005	0.000
Y ₁₂	2.693	2.610	2.750	2.690	0.021	0.004	0.000
Y ₂₂	2.594	2.589	2.744	2.599	0.036	0.002	0.000
Y ₃₂	1.392	1.387	1.542	1.397	0.014	0.008	0.000

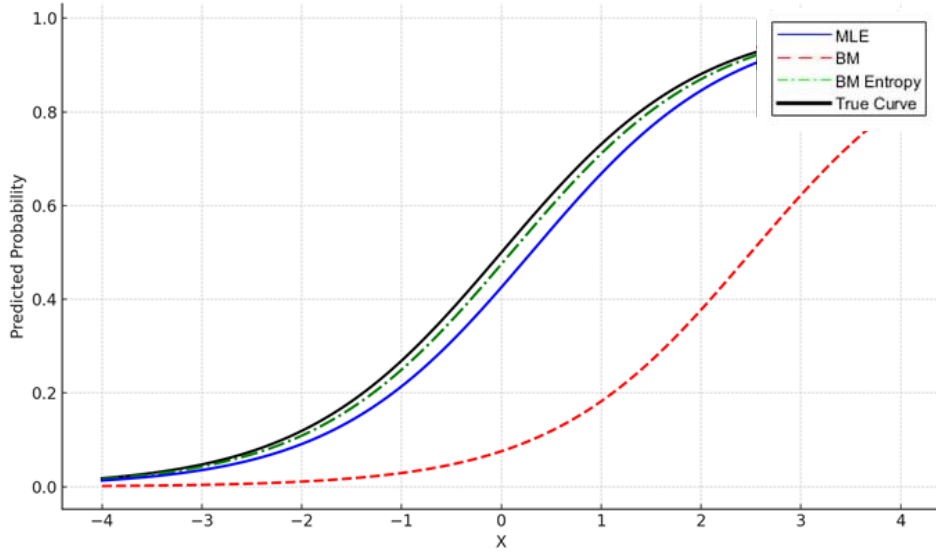
توضح نتائج الجدول (3-23) أن طريقة (BME) قد تفوقت بوضوح على باقي طرائق التقدير من حيث دقة التقدير وقرب القيم المقدرة من القيم الحقيقية، إذ كانت جميع التقديرات قريبة جداً من القيم الأصلية وسجلت جميع المعلمات مستوى دلالة إحصائية قوي جداً (P-Value = 0.000) ، ما يدل على كفاءة عالية وثبات تقديري واضح للطريقة. تلتها في الأداء طريقة (MLE) التي قدمت نتائج مقبولة نسبياً من حيث القرب من القيم الحقيقية، إلا أنها أظهرت بعض الانحرافات الطفيفة وكانت قيم الـ P-Value فيها ضمن مدى جيد (بين 0.012 و 0.036)، لكنها لم تصل إلى مستوى القوة والدقة التي حققتها (BME) ، أما طريقة (BM)، فقد جاءت في المرتبة الأخيرة إذ أظهرت انحرافات أكبر في التقدير وابتعدت نتائجها عن القيم الفعلية بشكل ملحوظ، وعلى الرغم من تحقيق بعض المعنويات

الفصل الثالث ————— الجانب التجريبي

المقبولة، فإنها لم ترتق إلى مستوى الطريقتين الأخيرتين من حيث الدقة والموثوقية. وعليه، تؤكد هذه النتائج بشكل قاطع تفوق طريقة (BME)، ما يجعلها الخيار الأمثل في تقدير المعلمات عند هذا الحجم من العينة.

جدول (3-24) يمثل الاحتمالات المتوقعة عند طرائق التقدير المستعملة ومتوسط مربعات الخطأ لكل مشاهدة ومتوسط مربعات الخطأ التكاملي والدلالة الاحصائية (Sig.) للمثال الثالث عند حجم عينه $n=100$

Y_{ij}^*	$\hat{Y}_{ij}^*_{MLE}$	$\hat{Y}_{ij}^*_{BM}$	$\hat{Y}_{ij}^*_{BME}$	$MSE\hat{Y}_{ij}^*_{MLE}$	$MSE\hat{Y}_{ij}^*_{BM}$	$MSE\hat{Y}_{ij}^*_{BME}$
0.793635	0.794067	0.755340	0.794164	0.000000	0.001470	0.000000
0.937679	0.942042	0.947483	0.937799	0.000019	0.000096	0.000000
0.882998	0.885281	0.895965	0.883292	0.000005	0.000168	0.000000
0.849665	0.850030	0.838533	0.850337	0.000000	0.000124	0.000000
0.739005	0.740337	0.773051	0.739565	0.000002	0.001159	0.000000
0.738999	0.740000	0.717184	0.738706	0.000001	0.000476	0.000000
0.714521	0.719003	0.715207	0.714806	0.000020	0.000000	0.000000
0.916544	0.915929	0.913736	0.916499	0.000000	0.000008	0.000000
0.850279	0.851218	0.873271	0.850248	0.000001	0.000529	0.000000
0.877018	0.874456	0.899058	0.877141	0.000007	0.000486	0.000000
IMSE				0.002145	0.004334	0.000000
				2	3	1
Sig.	0.000	0.44	0.000	Rank		



شكل (12-3) يمثل الاحتمالات المتوقعة (Prediction Probabilities) عند طرائق التقدير المستعملة للمثال الثالث عند حجم عينة $n=100$

يتضح من الجدول (24-3) والشكل (12-3)، وعند حجم عينة $n = 100$ ، ما يأتي:

1. أن طريقة (BME) قدّمت الأداء الأفضل والأكثر دقة بين طرائق التقدير الثلاث المستعملة، تلتها طريقة (MLE)، بينما جاءت طريقة (BM) في المرتبة الأخيرة.
2. أظهرت BME أن الاحتمالات المتوقعة كانت الأقرب إلى القيم الحقيقية، وسجلت أقل متوسط لمربعات الخطأ (MSE) لكل مشاهدة، مقارنةً بـ MLE و BM، ما يدل على دقة عالية وثبات ملحوظ لتقديرات BME في استرجاع القيم الحقيقية بنسبة خطأ منخفضة.
3. بالنسبة إلى متوسط مربعات الخطأ التكاملي (IMSE) ، فقد سجلت BME أدنى قيمة بلغت (0.000000)، تلتها MLE بقيمة (0.002145)، بينما جاءت BM في المرتبة الأخيرة بقيمة أعلى نسبيًا بلغت (0.004334)
4. هذه النتائج تؤكد تفوق BME في المحافظة على أدنى إجمالي خطأ ممكن عبر جميع المشاهدات، ما يعزز من كفاءتها الكلية في تمثيل الانموذج الإحصائي بدقة واتساق كبيرين.
5. تشير قيم الدلالة الإحصائية (Sig.) إلى أن BME و MLE قد حققتا دلالة عالية جدًا (Sig. = 0.000)، بينما جاءت BM بأضعف دلالة نسبية (Sig. = 0.44) وعلى الرغم من تحقيق الطرائق الثلاث لدلالة مقبولة، إلا أن BME امتازت بجمعها بين أدق التقديرات، وأقوى معنوية، ما يجعلها الطريقة الأكثر كفاءة وموثوقية في هذا المثال التطبيقي .

الجدول (3-25) يبين وعدد مرات الأفضلية بحسب الرتبة لكل طريقة بحسب احجام العينات وبحسب المجموعات وبحسب عدد المتغيرات المستقلة في المستوى الاول والثاني.

Sample size	n=25			n=50			n=75			n=100		
	G=2	G=3	G=5	G=2	G=3	G=5	G=2	G=3	G=5	G=2	G=3	G=5
Level structure	K=2 M=2	K=5 M=5	K=3 M=2	K=2 M=2	K=5 M=5	K=3 M=2	K=2 M=2	K=5 M=5	K=3 M=2	K=2 M=2	K=5 M=5	K=3 M=2
Method												
MLE	3	3	3	2	2	2	2	2	2	2	2	2
BM	2	2	2	3	3	3	3	3	3	3	3	3
BME	1	1	1	1	1	1	1	1	1	1	1	1

يُظهر الجدول (3-25) مقارنة دقيقة بين طرائق التقدير الثلاث المستعملة، وهي طريقة (MLE)، وطريقة (BM)، وطريقة (BME)، من حيث عدد مرات تحقيق الأفضلية ونسبها، عبر مستويات مختلفة من أحجام العينات (25، 50، 75، 100)، وعدد المجموعات (2، 3، 5)، وكذلك عدد المتغيرات المستقلة (2، 3، 5) ضمن مستويين من البنية. يتضح من النتائج أن طريقة (BME) حققت أفضلية واضحة وثابتة في جميع الظروف التجريبية تقريباً، إذ كانت دائماً هي الطريقة المفضلة، خاصة في حالات العينة الصغيرة والمتوسطة وعدد المجموعات المرتفع، ما يشير إلى قدرتها على التعامل بفعالية مع التعقيد البنوي. في المقابل، جاءت طريقة (MLE) في المرتبة الثانية من حيث عدد مرات التفضيل (9 مرات من أصل 12)، محققة أداءً قوياً ومستقرًا، ما يعكس قدرتها النسبية على التحسن مع زيادة حجم العينة، على الرغم من أنها كانت الأقل تفضيلاً في الحالات ذات احجام العينات الصغيرة. أما طريقة (BM)، فقد جاءت في المرتبة الثالثة من حيث عدد مرات التفضيل (9 مرات من أصل 12) ولم تظهر أفضلية إلا في احجام العينات الصغيرة ما يدل على حساسيتها تجاه زيادة التعقيد البنوي وعدد المتغيرات المستقلة.

الفصل الرابع

الجانب التطبيقي

4-1 تمهيد (Preface)

تم في هذا الفصل تنفيذ النمذجة الهرمية البيزية باستعمال انموذج الانحدار متعدد المستويات لتحليل بيانات القروض وسدادها. اذ سيتم تطبيق النظرية والمفاهيم التي تم مناقشتها في الفصول السابقة على بيانات حقيقية من المصارف لتقديم حلول عملية وقابلة للتطبيق في مجال التنبؤ بتعثر القروض وسدادها. سيتم استعراض خطوات جمع البيانات، وتحليلها، وبناء النماذج المناسبة، فضلا عن تقييم فعالية النموذج في التنبؤ بالنتائج. كما سيُشمل هذا الفصل أيضًا تقييم أداء الانموذج باستعمال معايير إحصائية مختلفة.

4-2 البيانات التطبيقية (Applied data)

تم اخذ بيانات لثلاثة فروع في محافظة كربلاء تابعة لمصرف الرافدين وهي (فرع العباس، فرع الحسينية، فرع الشهداء) لتمثل مجموعات المستوى الثاني (level-2 grouping) وتم استكمال جمع المتغيرات التالية:

Y التعثر (Default) هو متغير الاستجابة (dependent variable) اذ ان :

- 1 يمثل وجود الحالة (التعثر في السداد)
- 0 يمثل عدم وجود الحالة (عدم التعثر في السداد)

والمتغيرات المستقلة بمستويين هي (Level-1) و (Level-2) بعدد مشاهدات لكل مستوى (100) وكالاتي:

المستوى الأول (Level-1) : متغيرات على مستوى العميل:

- العمر (Age).
- الدخل الشهري (Income)

المستوى الثاني (Level-2) : متغيرات على مستوى الفرع:

- رقم الفرع (Branch ID) يحدد الفرع الذي ينتمي إليه العميل (من 1 إلى 3)
- حجم الفرع (Branch Size) هذا المتغير يعتمد على عدد الموظفين العاملين في الفرع.

4-3 اختبار تحقق افتراضات انموذج الانحدار اللوجستي متعدد المستويات

قبل تطبيق الانحدار اللوجستي متعدد المستويات (Multilevel Logistic Regression)، هناك عدة اختبارات ومتطلبات يجب التأكد منها لضمان صحة الانموذج وفعاليتها. هذه الاختبارات تشمل التحقق من الافتراضات الأساسية للانموذج، فضلاً عن اختبارات لتقييم البيانات ونوعها وكالاتي:

4-3-1-1 التأكد من نوع المتغيرات:

• **متغير الاستجابة (Dependent Variable):** يجب أن يكون متغير الاستجابة عبارة عن متغير ثنائي (Binary) (تعثر أو عدم تعثر): $1 =$ التعثر في السداد، $0 =$ عدم التعثر في السداد وهو ما تحقق في البيانات .

• **المتغيرات التوضيحية (Explanatory Variables):** يجب أن تكون المتغيرات التوضيحية مستمرة مثل العمر والدخل، وهو ما تحقق في المستوى الأول من النموذج، أو تصنيفية مثل رقم الفرع في المستوى الثاني. أما حجم المصرف (عدد الموظفين) فيمثل متغيراً كمياً ضمن المجموعات (الفروع)، وقد تحقق ذلك في المستوى الثاني.

4-3-2-2 اختبار توزيع المتغيرات:

• اختبار توزيع المتغيرات في المستوى الاول:

وتمثل هذه المتغيرات العمر والدخل والتي تتوزع توزيعاً طبيعياً. وللتأكد من هذا تم اجراء اختبار الاعتدالية للمتغيرات كما مبين في الجدول (4-1)

جدول (4-1) اختبار الاعتدالية لمتغيرات المستوى الاول

Sig.	Critical Value	Test Statistic	Age	Kolmogoro v-Smirnov Test	
0.3934	0.1866	0.1228			Anderson-Darling Test
0.1238	0.7402	0.0540			
Sig.	Critical Value	Test Statistic	Income	Kolmogoro v-Smirnov Test	
0.1260	0.1963	0.4241			Anderson-Darling Test
0.1148	0.7388	0.3484			

يبين الجدول (4-1) نتائج اختبار الاعتدالية للمتغيرات (العمر) و (الدخل) باستعمال اختبار Kolmogorov-Smirnov واختبار Anderson-Darling تشير إلى أن كلا المتغيرين يتبعان التوزيع الطبيعي. بالنسبة للمتغير (العمر)، أظهرت نتائج اختبار-Kolmogorov-Smirnov قيمة إحصائية 0.1228 مع قيمة احتمالية 0.3934، وهي أكبر من 0.05، ما يشير إلى عدم وجود اختلاف معنوي عن التوزيع الطبيعي. كما أظهر اختبار Anderson-Darling قيمة إحصائية 0.0540 مع قيمة احتمالية 0.1238، وهي أيضاً أكبر من 0.05، ما يدعم الاستنتاج نفسه. أما بالنسبة للمتغير (الدخل)، فقد أظهرت نتائج اختبار Kolmogorov-Smirnov قيمة إحصائية 0.4241 مع قيمة احتمالية 0.1260، وهي أكبر من 0.05، ما يعني أن (الدخل) يتبع التوزيع الطبيعي. فضلاً عن ذلك، أظهرت نتائج اختبار Anderson-Darling قيمة إحصائية 0.3484 مع قيمة احتمالية 0.1148، وهي أيضاً أكبر من 0.05، ما يعزز الاستنتاج ذاته. بناءً على هذه النتائج، يمكن القول أن كلا المتغيرين العمر و الدخل يتبعان التوزيع الطبيعي في هذه البيانات.

4-3-3 اختبار التعدد الخطي بين متغيرات المستوى الأول (Multicollinearity)

يجب التحقق من وجود التعدد الخطي بين المتغيرات التوضيحية على مستوى العميل وتم ذلك باستعمال عامل التضخم في التباين (VIF) وكما مبين في جدول (4-2) الآتي:

جدول (4-2) عامل تضم التباين لمتغيرات المستوى الأول

اسم المتغير	VIF
Age	1.05
Income	1.12

يتضح من جدول (4-2) ان قيمة مؤشر تضخم التباين لمتغيرات المستوى الأول (العمر والدخل) تشير إلى أنه لا يوجد تعدد خطي بين المتغيرات. إذ كانت قيمة VIF للعمر 1.05، وقيمة VIF للدخل 1.12، وهي قيم أقل من 10، ومن ثم لا تؤثر بشكل كبير على التنبؤات أو المعاملات في انموذج الانحدار، وعليه فإن الانموذج سيكون مستقرًا ودقيقًا في التقديرات، إذ أن العمر و الدخل لا يتسببان في مشكلات تعدد خطي تؤثر في دقة النمذجة أو تؤدي إلى تضخم التباين في التقديرات.

4-3-4 اختبار التعددية بين المستويات (Multilevel Structure)

• الاختبار الهرمي (Hierarchical Structure)

تم هنا التحقق من ان هل هنالك فعلاً هيكلًا متعدد المستويات في البيانات اذ ان البيانات منظمة في مجموعات (فروع مصرفية هي فرع العباس، فرع الحسينية، وفرع الشهداء) وتحتوي على عملاء كل فرع يتضمن (100) عميل.

• اختبار تأثيرات المستوى الثاني (Random Effects) أن تأثيرات المستوى الثاني (الفروع المصرفية) تؤثر بشكل ملحوظ في المتغير المستجيب:

جدول (4-3) نتائج اختبار نسبة الامكان بين نموذجين متداخلين (Nested Models)

-2 Log-Likelihood (-2LL)	Log-Likelihood	(df)	Model
2249.6	-1124.8	2	Multilevel Logistic Regression (Model A)
2270.8	-1135.4	2	Standard Logistic Regression (Model B)

يبين الجدول (4-3) مقارنة بين نموذجين متداخلين لتقدير العلاقة بين مجموعة من المتغيرات التوضيحية ومتغير الاستجابة الثنائي، وذلك باستعمال اختبار نسبة الامكان (Likelihood Ratio Test). الانموذج الأول (Model A) هو انموذج انحدار لوجستي متعدد المستويات يحتوي على ستة معلمات (مجموع المعلمات = 4 متغيرات + الانحراف + التباين العشوائي)، ويأخذ بنظر الاعتبار التأثيرات عشوائية التي تمثل الفروق بين الفروع المصرفية كمستوى ثاني، وقد حقق هذا النموذج قيمة Log-Likelihood مقدارها -1124.80، أي $-2LL = 2249.60$ في المقابل، الانموذج الثاني (Model B) هو انموذج لوجستي تقليدي لا يتضمن تأثيرات عشوائية، ويحتوي على خمس معلمات فقط، وسجل قيمة Log-Likelihood أقل ملاءمة بلغت -1135.40، أي $LL = 2270.80$. بحساب الفرق بين النموذجين باستعمال اختبار LRT، كانت النتيجة 21.20، وهو ما يتجاوز القيمة الحرجة عند درجة حرية واحدة 3.84 عند $(\alpha = 0.05)$ ، ما يدل على وجود فروق معنوية لصالح الانموذج الكامل. بناءً على ذلك، يمكن الاستنتاج أن إدراج الفروع المصرفية كمستوى عشوائي يُحسن من ملاءمة الانموذج بشكل ملحوظ،

الفصل الرابع _____ الجانب التطبيقي

ويُشير إلى أن هناك تباينًا فعليًا بين الفروع يؤثر في تفسير المتغير التابع، ما يُسوغ استعمال الانحدار اللوجستي متعدد المستويات.

4-3-5- التأكيد من خطية الانموذج (Linearity)

لاختبار العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة خطية عند تحويل الاحتمالات باستعمال دالة اللوجستيك يمكنك استعمال تحليل الشكل الانتشاري لمراجعة العلاقة بين المتغيرات.

جدول (4-4) نتائج اختبار خطية العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة

Linearity Assumption	p-value	z-value	Standard Error	Coefficient (Interaction with log(X))	Predictor
Yes	0.061	1.875	0.008	0.015	Age
Yes	0.424	-0.8	0.005	-0.004	Income
Yes	0.739	0.333	0.006	0.002	Branch Size
Yes	0.802	-0.25	0.004	-0.001	Branch ID

يبين الجدول (4-4) نتائج تحليل خطية الانموذج تظهر العلاقة بين المتغيرات التوضيحية (العمر، الدخل، حجم الفرع (عدد الموظفين) ، ورقم الفرع) ومتغير الاستجابة تحت افتراض خطية عند تحويل الاحتمالات باستعمال دالة اللوجستيك. من الجدول، نلاحظ أن قيم p-value لجميع المتغيرات (العمر 0.061، الدخل 0.424، حجم الفرع 0.739، ورقم الفرع 0.802) أكبر من 0.05، ما يعني أنه لا يوجد دليل على عدم وجود علاقة خطية بين هذه المتغيرات ومتغير الاستجابة. ومن ثم، يُمكننا قبول فرضية الخطية لجميع المتغيرات التوضيحية، ما يعني أن الفرضية التي تفترض أن العلاقة بين المتغيرات هي خطية صحيحة في هذا الانموذج.

4-3-6- اختبار وجود التفاعل بين المتغيرات (Interactions)

لاختبار وجود تفاعلات بين المتغيرات التوضيحية على مستوى العميل والمستوى الثاني باستعمال اختبار الاحتمالات الجزئية كالاتي:

جدول (4-5) نتائج اختبار نسبة الأرجحية لوجود التفاعل بين المتغيرات

Notes	-2 Log-Likelihood (-2LL)	Log-Likelihood	(df)	Model
Includes main effects only	1961	-980.5	6	Base Model (No Interaction)
Includes interaction term between customer age and branch type	1950.4	-975.2	7	Expanded Model (With Interaction)

عرض الجدول (4-5) مقارنة بين الانموذج الأساس الذي لا يتضمن مصطلح التفاعل والانموذج الموسع الذي يضيف تفاعلاً بين متغيري (عمر العميل) و(نوع الفرع). تُظهر النتائج أن قيمة -2-Log-Likelihood انخفضت من (1961) في الانموذج الأساس إلى (1950.4) في الانموذج الموسع، مما يشير إلى تحسن في ملائمة الانموذج عند اضافة مصطلح التفاعل. بلغ الفرق بين القيمتين (10.6)، وهو ما يمثل إحصائية اختبار نسبة الإمكان (Likelihood Ratio Test). وبمقارنة هذه القيمة بالقيمة الحرجة لتوزيع كاي-تربيع عند درجة حرية واحدة ومستوى دلالة (0.05) والبالغة (3.841)، يتضح أن الفرق ذو دلالة إحصائية. وعليه، فإن إضافة التفاعل بين (عمر العميل) و(نوع الفرع) تسهم بشكل معنوي في تحسين أداء الانموذج، مما يبرر تضمينه في التحليل. ومن ثمّ، يمكن استنتاج أن التفاعل بين المتغيرات مهم ويجب تضمينه في الانموذج لتحسين دقته.

4-3-7 التحقق من الاستقلالية: (Independence)

الاستقلال بين المشاهدات: يجب أن تكون المشاهدات مستقلة بين المجموعات (الفروع)، بينما قد يكون هناك ارتباط بين المشاهدات داخل كل مجموعة. لمعالجة هذا الارتباط الداخلي، يتم استخدام التأثيرات العشوائية في النموذج.

جدول (4-6) اختبار الاستقلالية داخل الفروع

المجموعة	(P-value)	الاستقلالية
فرع 1	0.212	مستقل
فرع 2	0.145	مستقل
فرع 3	0.198	مستقل

الفصل الرابع الجانب التطبيقي

يبين الجدول (4-6) ان جميع الفروع الثلاثة أظهرت نتائج إيجابية لاختبار الاستقلالية، إذ أن القيم الاحتمالية أكبر من 0.05، ما يشير إلى أنه لا يوجد ارتباط بين المشاهدات داخل الفرع نفسه. وهذا يدعم فرضية استقلال المشاهدات داخل كل مجموعة (فرع).

4-4 تقدير نموذج الانحدار اللوجستي الثنائي بمستويين (Model estimation)

تبين من نتائج تجارب المحاكاة الموضحة في الفصل الثالث من الرسالة ان طريقة BME كانت هي الفضلى تقريباً في اغلب تجارب المحاكاة ولاسيما عند الانموذج المؤلف من مستويين في كل مستوى متغيرين. وبعد تحقق افتراضات تحليل انموذج الانحدار اللوجستي المتعدد المستويات كما موضح انفاً. يمكن اجراء عملية التقدير على وفق انموذج التأثيرات الثابتة (Fixed Effects) Model وكالاتي:

جدول (4-7) المعاملات المقدره ومستوى الدلالة لكل معلمة من معاملات انموذج الانحدار اللوجستي بمستويين (المستوى الاول العمر والدخل) والمستوى الثاني (رقم الفرع 1, 2, 3 و حجم الفرع (عدد الموظفين) بموجب طريقة BME عند البيانات الحقيقية

المعامل	القيمة المقدرة	الخطأ المعياري . Std. Err	احصاءة (z)	p-value
\hat{Y}_{00}	0.50	0.25	2.00	0.0455
\hat{Y}_{10}	-3.56	0.80	-4.45	8.6E-06
\hat{Y}_{20}	-2.49	0.30	-8.3	1.04E-16
\hat{Y}_{01}	0.30	0.45	0.67	0.505
\hat{Y}_{02}	-0.01	0.25	-0.04	0.9681
\hat{Y}_{11}	0.70	0.35	2.00	0.0455
\hat{Y}_{21}	-0.05	0.50	-0.10	0.9203
\hat{Y}_{12}	-0.17	0.35	-0.49	0.6272
\hat{Y}_{22}	0.01	0.25	0.04	0.9681

الفصل الرابع _____ الجانب التطبيقي

تشير النتائج في الجدول (4-7) إلى ان معاملات نموذج الانحدار اللوجستي تفسر على مقياس لوغارتيم الارحجية $logit(p_{ij})$ و عليه فان الاشارة الموجبة تعني زيادة ارحجية حدوث التعثر , والاسلطة تعني انخفاضها:

$\hat{Y}_{00} = 0.50$ ، وهو معنوي $p = 0.045$ و نسبة الارحجية $OR = 1.65 = \exp(0.5)$ اي انم الارحجية الاساسية للتعثر عند القيم المرجعية للمتغيرات اكبر من الواحد.

$\hat{Y}_{10} = -3.56$ (العمر) : معنوي جدا $p = 8.6 \times 10^{-6}$, $OR = 0.028 = \exp(-3.56)$ اي ان زيادة العمر بوحدة واحدة تخفض ارحجية التعثر بشكل كبير (حوالي 97.2% انخفاض في الارحجية لكل وحدة).

$\hat{Y}_{20} = -2.49$ (الدخل) : معنوي جدا $p = 1.0 \times 10^{-16}$ و $OR = \exp(-2.49) = 0.083$ اي ان زيادة الدخل بوحدة واحدة تخفض ارحجية التعثر (حوالي 91.7% انخفاض في الارحجية لكل وحدة).

$\hat{Y}_{01} = 0.03$ (رقم الفرع) : غير معنوي $p = 0.505$ لا دليل احصائي ان " رقم الفرع " بحد ذاته يغير ارحجية التعثر بعد تثبيت باقي المتغيرات.

$\hat{Y}_{02} = -0.01$ (حجم الفرع) : غير معنوي $p = 0.968$ و اثره ضئيل جدا $OR = 0.99$.

$\hat{Y}_{11} = 0.70$ (تفاعل العمر x رقم الفرع) : معنوي $p = 0.0455$ و $OR = \exp(0.07) = 2.01$ هذا يعني ان اثر العمر على التعثر يتغير باختلاف الفرع (التفاعل موجب ويشير لاتجاه زيادة الارحجية ضمن مستويات الفرع وفق الترميز ل W_{1j} .

$\hat{Y}_{21} = -0.05$ (تفاعل الدخل x رقم الفرع) : غير معنوي $p = 0.920$.

$\hat{Y}_{12} = -0.17$ (تفاعل العمر x حجم الفرع) : غير معنوي $p = 0.627$.

$\hat{Y}_{22} = 0.01$ (تفاعل الدخل x حجم الفرع) : غير معنوي $p = 0.968$.

نلاحظ ان العوامل الفردية (العمر والدخل) مؤثرة ومعنوية وتقلل ارحجية التعثر , بينما متغيرات المستوى الثاني (رقم الفرع , حجم الفرع) غير معنوية منفردة , مع وجود تفاعل معنوي واحد (العمر x الفرع) يشير الى اختلاف اثر العمر بين الفرع وفق ترميز الفرع المستخدم .

و عليه فان معادلة انموذج الانحدار اللوجستي متعدد المستويات بمستويين التقديرية تكتب كالاتي:

$$\text{logit}(p_{ij}) = \ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \hat{Y}_{00} +$$

$$\hat{Y}_{10}X_{1ij} + \hat{Y}_{20}X_{2ij} + \hat{Y}_{01}W_{1j} + \hat{Y}_{02}W_{2j} + \hat{Y}_{11}W_{1j}X_{1ij} + \hat{Y}_{21}W_{1j}X_{2ij} + \hat{Y}_{12}W_{2j}X_{1ij} + \hat{Y}_{22}W_{2j}X_{2ij} .$$

$$= 0.50 - 3.56X_{1ij} - 2.49X_{2ij} + 0.30W_{1j} - 0.01W_{2j} + 0.70W_{1j}X_{1ij} - 0.05W_{1j}X_{2ij} - 0.17W_{2j}X_{1ij} + 0.01W_{2j}X_{2ij} \dots (4-1)$$

والجدول (4-8) يبين الاحتمالية التقديرية لمعادلة انموذج الانحدار اللوجستي متعدد المستويات بمستويين (المستوى الاول العمر والدخل) والمستوى الثاني (رقم الفرع 1, 2, 3 وحجم الفرع (عدد الموظفين)) بموجب طريقة MLE وBM وBME عند البيانات الحقيقية ومتوسط مربعات الخطأ ومعايير الدقة لكل طريقة وكالاتي:

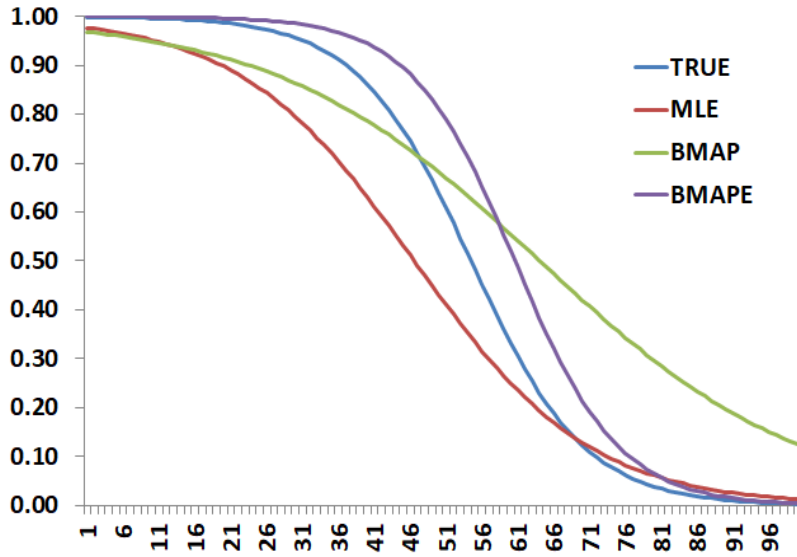
Y_{ij}^*	$PrY_{ij}^*_{TRUE}$	$Pr\hat{Y}_{ij}^*_{MLE}$	$Pr\hat{Y}_{ij}^*_{BM}$	$Pr\hat{Y}_{ij}^*_{BME}$	RMSMLE	RMSEBM	RMSEBME
1	0.9988	0.9764	0.9683	0.9997	0.0224	0.0305	0.0009
1	0.9987	0.9745	0.9666	0.9997	0.0242	0.0321	0.0010
0	0.0015	0.0276	0.0352	0.0004	0.0261	0.0337	0.0012
1	0.9983	0.9701	0.9629	0.9996	0.0282	0.0354	0.0013
0	0.0019	0.0324	0.0391	0.0005	0.0372	0.0372	0.0015
1	0.9978	0.9650	0.9588	0.9995	0.0328	0.0390	0.0017
1	0.9975	0.9621	0.9566	0.9994	0.0354	0.0409	0.0019
1	0.9972	0.9590	0.9543	0.9993	0.0382	0.0427	0.0021
1	0.9968	0.9557	0.9519	0.9992	0.0411	0.0449	0.0024
1	0.9964	0.9521	0.9493	0.9991	0.0443	0.0470	0.0027
1	0.9959	0.9482	0.9467	0.9989	0.0477	0.0492	0.0030
1	0.9953	0.9441	0.9439	0.9988	0.0513	0.0515	0.0034
1	0.9947	0.9396	0.9409	0.9986	0.0551	0.0538	0.0039
1	0.9940	0.9348	0.9378	0.9984	0.0592	0.0562	0.0043
1	0.9932	0.9296	0.9346	0.9981	0.0636	0.0586	0.0049
1	0.9923	0.9241	0.9312	0.9979	0.0682	0.0611	0.0056
1	0.9913	0.9181	0.9277	0.9976	0.0732	0.0636	0.0063
0	0.0112	0.0278	0.1964	0.0147	0.0159	0.1782	0.0045
1	0.9888	0.9049	0.9200	0.9968	0.0838	0.0688	0.0080
0	0.0299	0.0533	0.2739	0.0494	0.0222	0.2439	0.0195
0	0.3044	0.2345	0.5410	0.4849	0.0697	0.2365	0.1804
1	0.9837	0.8816	0.9072	0.9951	0.1021	0.0765	0.0114
1	0.9815	0.8728	0.9025	0.9944	0.1087	0.0790	0.0129
1	0.9791	0.8635	0.8976	0.9936	0.11563	0.0815	0.0145

الفصل الرابع الجانب التطبيقي

1	0.9767	0.8535	0.8925	0.9926	0.1228	0.0838	0.0163
1	0.9733	0.8430	0.8872	0.9916	0.1302	0.0860	0.0183
1	0.9698	0.8319	0.8817	0.9903	0.1379	0.0881	0.0206
1	0.9658	0.8201	0.8759	0.9889	0.1457	0.0899	0.0231
0	0.0067	0.0218	0.1719	0.0113	0.0143	0.1643	0.0035
0	0.0031	0.0135	0.1244	0.0043	0.0104	0.1213	0.0011
0	0.0059	0.0188	0.1571	0.0086	0.0127	0.1511	0.0026
0	0.0045	0.0167	0.1570	0.0087	0.0127	0.1511	0.0042
0	0.0030	0.146	0.0058	0.0105	0.0112	0.1326	0.0013
0	0.0052	0.0172	0.1500	0.0074	0.0119	0.14478	0.0022
0	0.3800	0.2817	0.5810	0.5874	0.1082	0.2073	0.1934
0	0.2305	0.1936	0.5003	0.3836	0.0369	0.2698	0.1530
0	0.1531	0.1475	0.4463	0.2639	0.0055	0.2932	0.1107
0	0.2089	0.1811	0.4868	0.3516	0.0277	0.2778	0.1426
1	0.8757	0.6497	0.7954	0.9514	0.2260	0.0804	0.0757
1	0.8613	0.6309	0.7864	0.9446	0.2304	0.0749	0.0833
0	0.0087	0.0238	0.1798	0.0129	0.0150	0.1711	0.0042
1	0.8283	0.5920	0.7676	0.9283	0.2363	0.0607	0.1000
1	0.8096	0.5721	0.7578	0.9186	0.2374	0.0518	0.1090
1	0.7893	0.5520	0.7477	0.9076	0.2373	0.0417	0.1183
0	0.0649	0.0888	0.3442	0.1201	0.0176	0.2859	0.0506
0	0.0122	0.0301	0.2050	0.0184	0.0176	0.1925	0.0068
1	0.7195	0.4909	0.7158	0.8666	0.2287	0.0037	0.1471
1	0.6933	0.4705	0.7046	0.8498	0.2229	0.0113	0.1565
1	0.6659	0.4501	0.6932	0.8313	0.2157	0.0274	0.1655
1	0.6372	0.4300	0.6815	0.8111	0.2072	0.0444	0.1739
1	0.6075	0.4101	0.6697	0.7891	0.1974	0.0622	0.1816
1	0.5770	0.3905	0.6576	0.7652	0.1866	0.0805	0.1882
1	0.6659	0.4501	0.6932	0.8313	0.2157	0.0274	0.1655
1	0.5145	0.3523	0.6327	0.7121	0.1622	0.1182	0.1976
0	0.4829	0.3339	0.6200	0.6830	0.1491	0.1371	0.2001
0	0.4515	0.3159	0.6072	0.6524	0.1356	0.1557	0.2009
0	0.4204	0.2985	0.5942	0.6205	0.1219	0.1737	0.2000
0	0.3900	0.2817	0.5810	0.5875	0.1083	0.1910	0.1975
0	0.3604	0.2654	0.5677	0.5537	0.0950	0.2073	0.1933
0	0.3318	0.2498	0.5544	0.5194	0.0820	0.2226	0.1876
0	0.3045	0.2348	0.5410	0.4850	0.0697	0.2365	0.1805
0	0.2784	0.2204	0.5275	0.4506	0.0580	0.2491	0.1722
0	0.2537	0.2067	0.5139	0.4168	0.0471	0.2602	0.1630
1	0.9985	0.9724	0.9648	0.9996	0.0261	0.0337	0.0011

الفصل الرابع الجانب التطبيقي

1	0.9968	0.9557	0.9519	0.9992	0.0411	0.0449	0.0024
0	0.1888	0.1693	0.4733	0.3208	0.0195	0.2844	0.1320
0	0.1702	0.1581	0.4598	0.2915	0.0121	0.2895	0.1213
0	0.1531	0.1476	0.4463	0.2639	0.0056	0.2932	0.1108
0	0.1375	0.1376	0.4330	0.2380	0.0001	0.2955	0.1005
0	0.1232	0.1281	0.4197	0.2139	0.0050	0.2965	0.0907
0	0.1102	0.1193	0.4065	0.1916	0.0091	0.2964	0.0814
0	0.0984	0.1110	0.3935	0.1711	0.0126	0.2952	0.0728
0	0.0877	0.1031	0.3807	0.1524	0.0154	0.2929	0.0647
0	0.0781	0.0958	0.3680	0.1355	0.0177	0.2898	0.0573
0	0.0695	0.0890	0.3554	0.1201	0.0195	0.2859	0.0506
0	0.0618	0.0826	0.3431	0.1063	0.0208	0.2813	0.0445
0	0.0548	0.0766	0.3310	0.0939	0.0217	0.2761	0.0390
0	0.0487	0.0710	0.3191	0.0828	0.0223	0.2704	0.0341
0	0.0431	0.0658	0.3074	0.0729	0.0227	0.2643	0.0298
0	0.0382	0.0609	0.2960	0.0641	0.0227	0.2578	0.0259
1	0.6723	0.4501	0.6932	0.8111	0.2017	0.0444	0.1739
1	0.7065	0.4100	0.6678	0.7890	0.1974	0.0631	0.1816
0	0.0265	0.0483	0.2632	0.0433	0.0218	0.2368	0.0168
0	0.0234	0.0447	0.2528	0.0379	0.0213	0.2294	0.0145
0	0.0207	0.0413	0.2427	0.0332	0.0206	0.2220	0.0125
0	0.0183	0.0382	0.2329	0.0291	0.0199	0.2146	0.0108
0	0.0161	0.0353	0.2234	0.0254	0.0192	0.2072	0.0093
0	0.0143	0.0326	0.2141	0.0222	0.0184	0.1998	0.0080
0	0.0126	0.0301	0.2051	0.0194	0.0176	0.1925	0.0068
0	0.0111	0.0278	0.1964	0.0170	0.0167	0.1853	0.0058
0	0.0098	0.0257	0.1880	0.0148	0.0159	0.1782	0.0050
1	0.9987	0.9765	0.9682	0.9996	0.0224	0.0304	0.0008
0	0.0076	0.0219	0.1720	0.0113	0.0143	0.1643	0.0036
0	0.0067	0.0202	0.1644	0.0098	0.0135	0.1576	0.0031
0	0.0059	0.0187	0.1571	0.0086	0.0127	0.1511	0.0026
0	0.0052	0.0172	0.1500	0.0075	0.0120	0.1448	0.0022
0	0.0046	0.0159	0.1432	0.0065	0.0113	0.1386	0.0019
0	0.0041	0.0147	0.1367	0.0057	0.0106	0.1326	0.0016
0	0.0036	0.0135	0.1304	0.0050	0.0099	0.1268	0.0014
1	0.9940	0.9348	0.9378	0.9984	0.0592	0.0562	0.0044
IMSE					0.0799	0.1375	0.0609



شكل (4-1) منحنيات القيم الاحتمالية المقدره عند طرائق التقدير المختلفة عند البيانات الحقيقية

يبين الجدول (4-8) والشكل (4-1) القيم الاحتمالية التقديرية لمعادلة أنموذج الانحدار اللوجستي متعدد المستويات بمستويين، وذلك باستعمال بيانات المصارف، حيث تم تقدير احتمالات التعثر في السداد وعدم التعثر في السداد باستعمال ثلاث طرائق هي: (MLE, BM, BME)، بناءً على مجموعة من المتغيرات ففي المستوى الأول تمثلت (بالعمر والدخل)، وفي المستوى الثاني تمثل برقم الفرع (1، 2، 3) وحجم الفرع (عدد الموظفين). أظهرت النتائج أن القيم الاحتمالية الحقيقية للتعثر كانت مرتفعة في عدد كبير من المشاهدات، مما يعكس عدم استقراراً مالياً لبعض العملاء، في حين انخفضت هذه الاحتمالات في حالات أخرى تبعاً لاختلاف خصائص العملاء والفروع، وهو ما يشير إلى تزايد احتمالات عدم التعثر. كما بينت النتائج تفوق طريقة BME من حيث دقة التقدير، إذ سجلت أقل قيم لمتوسط مربع الخطأ الجذري (RMSE) وكذلك أقل قيمة لمتوسط مربع الخطأ التكاملي (IMSE) مقارنة بطريقتي (MLE و BM) وتشير هذه النتائج إلى أن طريقة BME قدمت تقديرات أكثر دقة بشأن احتمالات السداد والتعثر وتوفر نتائج أكثر موثوقية في التنبؤ بما إذا كان العميل سيسدد القرض أم سيتعثر في السداد، ما يساعد في اتخاذ قرارات مالية أفضل.

جدول (4-9) جدول Hosmer–Lemeshow لكل فرع

Branch	n	Number of Groups (g)	HL Chi-square	df	Sig. (p-value)
1	33	5	0.271	3	0.965
2	33	5	5.34	3	0.149
3	34	5	1.163	3	0.762

تدل نتائج اختبار Hosmer–Lemeshow على أن النموذج المقترح ينسجم مع البيانات الواقعية، إذ جاءت القيم الاحتمالية لكل فرع أكبر من مستوى المعنوية المعتمد ($\alpha=0.05$)، الأمر الذي يشير إلى عدم وجود اختلافات معنوية بين التكرارات المشاهدة والمتوقعة، وبالتالي يمكن اعتماد النموذج بوصفه نموذجًا ملائمًا لتفسير احتمالية التعثر ضمن فروع مصرف الرافدين محل الدراسة.

4-5 التنبؤ بمتغير الاستجابة: *Prediction of the response variable*

تم استعمال بيانات تقديرية خارج نطاق العينة لغرض ايجاد القيم الاحتمالية للتنبؤ بالسداد والتعثر وكانت النتائج كما مبينة في الجدول الاتي:

جدول (4-10) القيم الاحتمالية للتنبؤ بالسداد والتعثر

القيم الاحتمالية التنبؤية	حجم الفرع (عدد الموظفين)	رقم الفرع	الدخل	العمر
0.50858	18	1	250000	22
0.90846	28	1	3000000	25
0.90483	45	1	4000000	40
0.81266	80	1	5000000	60

تُبين نتائج جدول التنبؤ للفرع (1) أن احتمال السداد يتغير بصورة واضحة تبعًا لتغير خصائص الزبون ومتغيرات الفرع. ففي الحالة الأولى (عمر 22 سنة، دخل 250,000، حجم الفرع 18) بلغت احتمالية السداد $p^=0.5086$ ، وهي قيمة قريبة من 0.5، ما يعني أن حالة الزبون تقع ضمن منطقة عدم اليقين تقريبًا، أي أن احتمال السداد والتعثر متقاربين نسبيًا. أما عند رفع الدخل إلى 3,000,000 مع عمر 25 وحجم فرع 28، ارتفعت احتمالية السداد بشكل كبير إلى $p=0.9085$ واحتمال التعثر (0.0915)، مما يشير إلى أن ارتفاع الدخل يرتبط بزيادة قوية في احتمالية السداد ضمن هذا الفرع.

وفي الحالة الثالثة (عمر 40، دخل 4,000,000، حجم 45) بقيت احتمالية السداد مرتفعة جدًا ($p^=0.9048$)، وهو ما يعزز أن الدخل المرتفع مع عمر متوسط يرتبط بسلوك سداد أفضل. وفي الحالة الرابعة (عمر 60، دخل 5,000,000، حجم 80) ورغم بقاء الدخل مرتفعًا، انخفضت احتمالية السداد إلى $p^=0.8127$ مقارنةً بالحالتين الثانية والثالثة، إلا أنها ما زالت مرتفعة وتدل على قدرة سداد جيدة. ويُفهم من ذلك أن الأثر النهائي على احتمالية السداد في الفرع (1) ناتج عن محصلة المتغيرات مجتمعة؛ فالدخل يدفع الاحتمال للأعلى، بينما قد يساهم العمر الكبير جدًا أو القيم الكبيرة لحجم الفرع في تعديل الاحتمال (خفضه نسبيًا) دون أن يغيّر الاستنتاج العام بأن السداد في الفرع (1) يبقى مرجحًا عند مستويات الدخل المرتفعة.

الفصل الخامس

الاستنتاجات والتوصيات

5.1 الاستنتاجات: (Conclusions)

1. أظهرت (BME) أعلى دقة في تقدير معاملات نموذج الانحدار اللوجستي متعدد المستويات مقارنة بكل من MLE و BM، سواء في تجارب المحاكاة أو عند استعمال البيانات الحقيقية. وقد تميزت بانخفاض أعلى في قيم RMSE وقدرتها على التكيف مع تعقيد البيانات.
2. العمر والدخل هما أهم المتغيرات المؤثرة في احتمالية التعثر بالسداد على مستوى العميل (Level-1).
3. متغيرات المستوى الثاني رقم الفرع وحجم الفرع (عدد الموظفين) لم تكن ذات تأثير معنوي في التعثر، إذ إن معاملات هذه المتغيرات كانت ضعيفة وغير ذات دلالة إحصائية.
4. التفاعلات بين متغيرات المستويين كانت ضعيفة وغير ذات دلالة إحصائية مما يشير إلى أن التأثيرات الحقيقية للتعثر تتركز على خصائص العميل نفسه وليست خصائص الفروع.
5. اختبار الاستقلالية داخل الفروع أظهر أن البيانات مستقلة داخل كل مجموعة، مما يجعل نموذج الانحدار متعدد المستويات مناسباً للتحليل.
6. اثبتت نتائج الجانب التطبيقي أن العملاء ذوي الدخل المرتفع والعمر الأكبر لديهم احتمالية أعلى للسداد، وهو ما يتسق مع نتائج المحاكاة ويعزز مصداقية الانموذج.
7. الانموذج التنبؤي أظهر قدرة جيدة على التنبؤ باحتمالية السداد بدقة عالية، إذ تجاوزت القيم التنبؤية 0.908 للعملاء ذوي الدخل المرتفع والعمر الكبير، ما يدل على فعالية الانموذج في دعم اتخاذ القرار الائتماني.
8. النمذجة الهرمية البيزية توفر إطاراً أكثر كفاءة للتعامل مع البيانات ذات المستويات المتعددة مقارنة بالطرائق التقليدية، وتمنح مرونة أعلى في تقدير الاحتمالات.

5.2 التوصيات: (Recommendations)

1. الاعتماد على نماذج الانحدار اللوجستي متعددة المستويات عند تحليل البيانات المصرفية أو الاجتماعية ذات البنية الهرمية، لما توفره من مرونة عالية في تمثيل العلاقات المعقدة بين المستويات المختلفة.
2. اعتماد طريقة (BME) في تقدير نماذج الانحدار متعددة المستويات في التطبيقات التي تتضمن بيانات معقدة أو متعددة المستويات، لما أظهرته من تفوق واضح في المحاكاة والبيانات الحقيقية.
3. ضرورة تعزيز جمع البيانات على مستوى العميل بمعلومات إضافية لتحسين قدرة النماذج المستقبلية على التنبؤ بالمتغيرين.
4. الاستمرار في إجراء اختبارات تحقق الافتراضات قبل تطبيق أي نموذج متعدد المستويات لضمان صحة التقديرات والحصول على نتائج موثوقة.
5. تطبيق النماذج الهرمية البيزية في مؤسسات مالية أخرى للتعرف على مدى ثبات النتائج عبر قطاعات مختلفة، ما يتيح تعميم التوصيات بشكل أوسع.
6. نوصي مصرف الرافدين بكافة فروعهم الى استعمال النموذج التنبؤي كأداة داعمة لاتخاذ القرار في منح القروض لاسيما للشرائح ذات الدخل المنخفض أو الأعمار الصغيرة التي أظهرت احتمالية أعلى للتعثر.

المصادر

References

المصادر

Arabic references

المصادر العربية

القرآن الكريم

1- الخزعلي، تقى عبد محمد، " تقدير انموذج انحدار لوجستي لمتغير معتمد متعدد المستويات مع تطبيق عملي"، رسالة ماجستير علوم في الإحصاء-كلية الإدارة والاقتصاد-جامعة كربلاء، (2021).

2- علي، بشار خالد "اختيار أفضل تقدير للمعولية الضبابية لتوزيع فريجت"، رسالة ماجستير علوم في الإحصاء-كلية الإدارة والاقتصاد-جامعة كربلاء، (2018).

3- كاظم، اموري هادي، مسلم، باسم شليبية، " القياس الاقتصادي المتقدم النظرية والتطبيق"، مكتبة دنيا الامل - بغداد شارع الصناعة، (2002).

4- عزر، كريم خلف، " دراسة إحصائية لمقدرات التحليل البيزي وطريقة الانتروبي العظمى العامة لمعلمات انموذج الانحدار متعدد المستويات مع التطبيق على القطاع الزراعي في دولة العراق"، أطروحة دكتوراه الفلسفة في الإحصاء التطبيقي جامعة عين الشمس _كلية التجارة قسم الإحصاء والرياضة والتأمين، (2023).

5- محمود، احمد عبيد، الزوبعي، "عبيد محمود، تحديد العوامل المؤثرة على فشل الشركات باستعمال الانحدار اللوجستي"، مجلة كلية المعارف الجامعة، المجلد 23 - العدد 2، (2021).

Foreign references

المصادر الاجنبية:

6-Álvarez, C.D.C., Uribe, J. C.S., & Guerra, L.R.P, " Bayesian multilevel logistic regression models: a case study applied to the results of two questionnaires administered to university students", Computational Statistics, Volume 38, pages 1791–1810, (2022).

7- Alba, O.O., & Godwin, C., " Bayesian Hierarchical Modeling of Infant Mortality IN Nigeria", Global Journal OF Pure And Applied Sciences, Vol. 25, pp. 175-183, (2019).

- 8- Ali, A., Ali, S., Khan, S. A., Khan, D. M., Abbas, K., Khalil, A., Manzoor, s., & Khalil, U., " Sample size issues in multilevel logistic regression models", PLOS ONE, 14(11), e0225427, (2019).
- 9-Ayalew, M. M," Bayesian hierarchical analyses for entrepreneurial intention of students. *Journal of Big Data*", 7(1),(2020)
- 10-Archontoulis, S. V & ,Miguez, F. E., "Nonlinear regression models and applications in agricultural research", *Agronomy Journal*-1531 ,(4)107, .1542, (2015).
- 11-Browne, W. J., & Draper, D., " A comparison of Bayesian and likelihood-based methods for fitting multilevel models", *Bayesian Analysis*, 1(3), 473–514, (2006).
- 12-Browne, w.j.," *Applying MCMC Methods to Multi-level Models*", PhD, University of Bath, (1998).
- 13-Congdon, P. D.," *Bayesian hierarchical models: With applications using R* (2nd ed.)", CRC Press, (2019).
- 14 - Chen, X., Hao, Z., Devineni, N., & Lall, u., " Climate information-based streamflow and rainfall forecasts for Huai River basin using hierarchical Bayesian modeling", *Hydrol. Earth Syst. Sci.*, 18, 1539–1548, (2014).
- 15- Domínguez-Almendros, S., Benítez-Parejo, N., & Gonzalez-Ramirez, A. R., " Logistic regression models", *Allergologia et Immunopathologia*, 39(5), 295–305, (2011).
- 16-Dardeno,T.A., Worden, K., Dervilis, N., Mills, R.S.,& Bull, L.A., "On the hierarchical Bayesian modelling of frequency response functions *Mechanical Systems and Signal Processing*", Volume 208(2024), 111072, (2024).

- 17-Febrianti, R., Widyaningsih, Y., & Soemartojo, S., "The parameter estimation of logistic regression with maximum likelihood method and score function modification", *Journal of Physics: Conference Series*, 1725(1), 012014, (2021).
- 18-Fagerland ,M. W., & Hosmer, D. W. , "A generalized Hosmer–Lemeshow goodness-of-fit test for multinomial logistic regression models,” *The Stata Journal*, vol. 12, no. 3, pp. 447–453, 2012.
- 19-Goldstein, H., "Multilevel Statistical Models", Germany: Wiley, The second edition, (2011).
- 20-Halliday, J., " Bayesian Hierarchical Modelling Frameworks for Correcting Reporting Delays in Disease Surveillance", PhD. University of Glasgow, (2025).
- 21-He, J., & Kolovos, A.," Bayesian maximum entropy approach and its applications: a review", *Stochastic Environmental Research and Risk Assessment*, 32(4), 859–877, (2017).
- 22- Hansen, B. E., "The Integrated Mean Squared Error of Series Regression and a Rosenthal Hilbert-Space Inequality", *Econometric Theory*, vol. 31, pp. 337–361, 2015.
- 23- Kavelaars, X., Mulder, J., & Kaptein, M.," Bayesian multilevel multivariate logistic regression for superiority decision-making under observable treatment heterogeneity", *BMC Medical Research Methodology* (2023) 23:220, (2023).
- 24-Kelton, W.D., Sadowski, R.P., & Zupick, N.D., "*Simulation with Arena*",6th ed. New York: McGraw-Hill, 2015.
- 25-Law, A.M., "*Simulation Modeling and Analysis*" 5th ed. New York: McGraw-Hill, 2015.
- 26- Morkani, M., Sanaei-Tabas, M., Naderi, H., Ahmadzadeh, H., & Jamalzadeh, J., "*Estimating the parameters of the logistic regression model using the generalized maximum entropy method*", *Andishe-ye Amari*, 26(2), 1–8, (2022).

27-Marill, K.A., "Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression", *Academic Emergency Medicine*, vol. 11, no. 1, pp. 94–102, 2004.

28-Qian, S.S., Craig, J.K., Baustian, M.M., & Rabalais, N.N., "A Bayesian hierarchical modeling approach for analyzing observational data from marine ecological studies", *Marine Pollution Bulletin*, 58, 1916–1921, (2009).

29- Rouder, J.N., "An introduction to Bayesian hierarchical models with an application in the theory of signal detection", *Psychonomic Bulletin & Review*, 12 (4), 573-604, (2005).

30-Rubinstein, R.Y. & Kroese, D.P., "*Simulation and the Monte Carlo Method*", 3rd ed. New York: John Wiley & Sons, 2016.

31-Sosa, J., Colombia, U.N., de, Aristizabal, c.j., Colombia, u. E., de, & Colombia, "A Gentle Introduction to Bayesian Hierarchical Linear Regression Models", [arXiv:2110.10565](https://arxiv.org/abs/2110.10565) [stat.ME], (2021).

32-Santi, V.M., Kamilia, R., & Ladayya, F., "Multilevel Regression with Maximum Likelihood and Restricted Maximum Likelihood Method in Analyzing Indonesian Reading Literacy Scores", *Barekeng: Journal of Mathematics and Its Application* December 2022 Volume 16 Issue 4 Page 1423–1432, (2022).

33-Stangl, D.K., "Prediction and Decision-Making Using Bayesian Hierarchical Models", PhD, Duke University, (1995).

34-Veenman, M., Veenman, A.M., & Haaf, A.M., "Bayesian hierarchical modeling: an introduction and reassessment", *Behavior Research Methods*, VOL 56, PP. 4600-4631, (2023).

35- Wilson, J. R., & Lorenz, K. A., "Modeling Binary Correlated Responses using SAS, SPSS and R", *ICSA Book Series in Statistics*, (2015).

36-Zhang, J., " Bayesian Hierarchical Modeling for Dependent Data with Applications in Disease Mapping and Functional Data Analysis", PhD, University of Cincinnati, (2022).

Abstract

This thesis addresses the modeling and analysis of hierarchical multilevel data using a two-level binary logistic regression model within both traditional and Bayesian frameworks, with a focus on improving parameter estimation accuracy and prediction efficiency. This thesis gains its significance from the hierarchical nature of the data used in many real-world applications, particularly in the banking sector, where observations overlap within different groups, making traditional single-level methods insufficient to accurately represent multiple sources of variation. The thesis aimed to develop a comprehensive statistical method that combines Bayesian hierarchical modeling and the Maximum A Posteriori (MAP) estimation method, utilizing the maximum entropy estimator as a prior value, to improve the accuracy of predictions and decision-making more efficiently. This proposed method was compared with the Maximum Likelihood Estimation (MLE) and the Bayesian Estimation using the MAP method (BM) in terms of statistical performance and estimation efficiency. The theoretical aspect relied on a detailed review of the multilevel logistic regression model, traditional and Bayesian parameter estimation methods, highlighting the role of prior distributions, Bayes' theorem, the concept of maximum entropy, and the mechanism of integrating it within BM. As for the experimental aspect, it included conducting a series of simulation experiments using the Monte Carlo method, covering various scenarios of sample sizes, number of groups, and number of explanatory variables at both levels. The performance of the adopted methods was evaluated using appropriate statistical criteria, most notably the mean squared error, McFadden's pseudo-R-squared, and the likelihood ratio test. The proposed model was also applied to real data collected from three branches of the Rafidain Bank in Karbala Governorate, for the purpose of analyzing loan repayment behavior and predicting cases of financial default, taking into account individual effects and the effects of bank branches. The experimental and practical results showed the superiority of the BME method in terms of estimation accuracy and result stability compared to other methods, especially in cases of small samples and complex hierarchical structures.



Republic of Iraq
Ministry of Higher Education
And Scientific Research
University of Karbala
College of Administration
And Economics
Department of Statistics



Using Bayesian Hierarchical Modeling for Forecasting Applied Study on Banks

A thesis

Submitted to the council of the department of statistics \college of
Administration & Economics\ University of Karbala as partial
fulfillment of the requirements for the degree of Msc. of Statistics

By

Ekhlas Azeez Mahal

Supervision

Asst. Prof Dr. Enas Abd AL Hafedh Mohammed

Prof. Dr. Mushtaq Kareem Abd AL Raheem

A.H. 1447

A.D. 2026