الفصل العاشر

مخاطر الائتمان

سجل التصنيف الائتماني لشركة أثناء الأوقات العصيبة

كانت ليز كليبورن مصممة أزياء أمريكية، ولدت في مدينة بروكسل في 1929. وفي عام 1976، كانت واحدة من الشركاء المؤسسين لشركة ليز كليبورن، مما جعلها تدخل ضمن قائمة مجلة فورتشن لأعلى 500 شركة في عام 1986. كانت أول امرأة تتولى منصب المدير التنفيذي في قائمة مجلة فورتشن لأعلى 500 شركة، وكان لها تأثيرًا كبيرًا على الطريقة التي تُباع من خلالها الأزياء، فقد أصرت على جمع الملابس في مجموعتها معًا في المتجر عند عرضها بدلاً من وضع تنورات ليز كليبورن مع التنورات الأخرى في مكان منفصل عن القمصان. واستقالت كليبورن من الإدارة النشطة في عام 1989 وتوفت في عام 2007.

كانت مبيعات شركة ليز كليبورن السنوية 2 مليار دولار في بدايات تسعينيات القرن الماضي وتوسعت عبر عمليات استحواذ على مدار التسعينيات وبدايات الألفية الجديدة، فقد استحوذت على لاكي براند جينز ومكس وجوسي كوتور، ولكنها في النهاية بدأت المشاكل تتوالى عليها. ربما كانت عمليات الاستحواذ كثيرة للغاية ولكن كانت هناك أيضًا

صعوبات في إدارة العلاقات مع كبار شركات التجزئة مثل مايسيز وجي سي بيني. تولى ويليام ماكونب منصب المدير التنفيذي للشركة في 2006 وارتفعت أسهم الشركة في بدايات 2007 إلى 46.64 دولارًا، ولكن كانت هناك مشاكل واضحة بحاجة إلى معالجتها: مجموعة من العملاء المسنين والصورة السيئة المتراجعة للعلامة التجارية في الصحف. ثم ازدادت الأمور سوءًا زيادة كبيرة، حيث بدأت الشركة تسجل خسائر في الربع الأخير من عام 2007 واستمرت هكذا حتى عام 2012، وانخفض سعر السهم انخفاضًا جذريًا مسجلاً أقل من دولارين في 2009 عندما كانت الشركة تُسرِّح عمالاً وتغلق مراكز توزيع. وفي عام 2011، سجلت الشركة خسارة فادحة (بقيمة 172 مليار دولار) ولكن استمر سعر السهم في تعافيه البطيء وبلغ في 2013 (بينما أكتب هذه السطور) حوالي 19 دولارًا.

من جهة أخرى، تملك الشركة – ولمدة طويلة – تصنيفًا ائتمانيًا من شركة ستاندرد آند بورز. وعند النظر إلى تاريخ الشركة منذ 2000 سنرى أن تصنيف شركة ليز كليبورن الائتماني كان عند الدرجة الاستثمارية (أي BBB بمصطلحات ستاندرد آند بورز) حتى 3يونيو 2008، عندما انخفض تصنيفها إلى درجة BB، وهي درجة أدنى من درجة الاستثمار (المعروفة باسم الدرجة الرديئة). في ذلك الوقت، ذكرت ستاندرد آند بورز أن السبب وراء هذا التخفيض هو الارتفاع الهائل في مستويات الديون والتحديات في بيئة البيع بالتجزئة في هذا المجال. ثم انخفض تصنيف شركة ليز كليبورن الائتماني مجددًا في 17 أغسطس 2009 إلى درجة B ومرة أخرى في مارس 2010 إلى درجة CC. ويعني هذا التصنيف أن ستاندرد آند بورز ترى الشركة "ضعيفة للغاية" وهو أيضًا إشارة تحذير هائلة. تلقت الشركة الإهانة الأخيرة في 11 إبريل 2011 عندما تم تصنيف الشركة عند درجة العجز الانتقائي. وعندما تختار ستاندرد آند بورز تصنيف مقدم عطاء بغرض الحصول على إعادة تمويل "متعثرًا" ومساويًا للعجز الانتقائي فغالبًا ما يكون هذا قرارًا فنيًا ومغايرًا تمامًا عن التخلف العادي عن الديون. وفي الواقع، أعيد تصنيف الشركة في اليوم التالي مباشرة إلى درجة B. ومنذ ذلك الوقت والشركة (التي غيرت اسمها التجاري إلى مجموعة شركات فيفث آند باسيفيك) في وضع أفضل بكثير على الرغم من بقاء التحديات، فقد تم

بيع بعض العلامات التجارية لسداد الديون وركزت الشركة على علاماتها التجارية الثلاثة الأساسية: كايت سبايد وجوسي كوتور ولاكي براند.

تمثل التصنيفات الائتمانية بالنسبة لشركة مثل ليز كليبورن بحاجة إلى اقتراض المال أهمية جوهرية. والهدف من وجود وكالات التصنيف الائتماني هو تقديم مؤشر على احتمالية عدم قدرة الشركات على الوفاء بالتزاماتها مستحقة السداد. ولكن إلى أي مدى تختلف المعلومات التي يوفرها التصنيف الائتماني عن تقييم سوق الأسهم؟ وما مدى موثوقية التصنيفات.

1-10 مقدمة عن مخاطر الائتمان

في هذا الفصل، سوف نعطي مقدمة عن خطر الائتمان بما يشمل مناقشة حول الائتمان الاستهلاكي (وهو موضوع من المرجح أن يكون ذا صلة بالمديرين العاملين خارج القطاع المالي). قد نكون مهتمين للغاية بخطر الائتمان لأنها تؤثر علينا نحن الأفراد، فبما أن ظروفنا الخاصة وتاريخنا الائتماني سيؤثر في قدرتنا على أموال وفي معدلات الفائدة التي ندفعها.

وكما ناقشنا في الفصل الأول، يُقصد بخطر الائتمان احتمالية أن يصبح عقدًا قابلاً للتنفيذ القانوني منعدم القيمة (أو على الأقل خفض قيمته خفضًا جوهريًا، لأن الطرف الآخر تخلف عن السداد وخرج من السوق. مع الأفراد، قد يصبح دين مستحق السداد غير قابل للتحصيل حتى بدون أن يصبح الشخص المعني مفلسًا، ويمكن تصنيف هذا أيضًا على أنه خطر ائتمان.

ستقوم مناقشتنا في هذا الفصل بمعالجة جانبين من مخاطر الائتمان، الجانب الأول هو امتلاك خطر الائتمان خصائص محددة، أي إما نعم أو لا مثل أن يكون هناك تخلف في السداد أو لا يكون. لم يعد التحليل التفصيلي لخطوط التوزيع والنتائج المحتملة ذا صلة بالسوق الحالي. والجانب الثاني هو ارتباط خطر الائتمان بما يحدث بالطرف الآخر، وبالتالي فنحن نملك معلومات أقل عن أعمالنا. نحن بحاجة إلى بذل مجهود أكبر للحصول على أقصى استفادة من أي معلومات متوفرة لدينا. ينبغي علينا النظر فيما يحدث مع مرور الوقت (هل تنتقل الشركة إلى وضع أقل استقرارًا؟، وينبغي علينا الاستفادة من المؤشرات غير المباشرة المتوفرة لدينا (هل يُعد إصدار هذا الفرد 4 بطاقات ائتمان جديدة خلال الأشهر الستة الماضية إشارة سلبية؟).

صممت اتفاقية بازل 2 للبنوك. وسترتبط خطر الائتمان في هذا السياق بعدد كبير من القروض متنوعة الأشكال، سواء قروض الشركات أو القروض الشخصية. وتمثل الرهون العقارية أحد المكونات الكبيرة لإقراض الأفراد. في هذه الحالة، توفر قيمة العقار نفسه بعض الأمان للبنك ولكن قد يظل هناك خطر ائتمان كبير إذا انخفضت أسعار العقارات انخفاضًا كبيرًا بدرجة تجعل الدين المستحق أكبر من قيمة المنزل (كما حدث في الولايات المتحدة خلال الأزمة العقارية).

ومن منظور مؤسسي، يرتبط خطر الائتمان بالتصنيفات الائتمانية التي تقدمها واحدة من وكالات التصنيف الائتماني الكبيرة الثلاثة: ستاندرد آند بورز وموديز وفيتش. ويستحسن عند الدخول في عقد مع شركة – وخاصة عند إقراضها مالاً– منخفضة التصنيف الائتماني وبالتالي يزداد خطر تخلفها عن السداد التمهل والتفكر في الأمر مليًا. فمن الحكمة –على أقل تقدير– طلب فائدة أعلى على القروض المنوحة للشركات التي ينظر إلى تصنيفها الائتماني على أنه خطر. وعن طريق قصر الترتيبات التعاقدية مع الشركات منخفضة التصنيف الائتماني، سيستطيع المديرون تقليل خطر الائتمان الذي يتعرضون له.

تملك كل وكالة تصنيف ائتماني مصطلحاتها أكوادها الخاصة بها، ووفيما يلي أكواد ستاندرد آند بورز:

الدرجة الاستثمارية

AAA : أفضل المقترضين، يتمتعون بالثقة والاستقرار (العديد منهم من الحكومات).

AA : قدرة قوية للغاية على الوفاء بالالتزامات المالية ومخاطر أعلى قليلاً من التصنيف AAA.

A : قدرة قوية على الوفاء بالالتزامات المالية ولكن مع احتمالية التأثر بسهولة من الأوضاع الاقتصادية السيئة.

BBB : مقترض متوسط يملك قدرة معتدلة على الوفاء بالالتزامات ووضعه اللحظي مناسب الدرجة غير الاستثمارية.

BB : غير ضعيف على المدى القريب ولكنه يواجه مشاكل كبرى قائمة.

B : يملك حالياً القدرة على الوفاء بالالتزامات ولكنه ضعيفًا أمام الظروف المعاكسة.

CCC : ضعيف حاليًا ويعتمد على الأعمال التفضيلية والظروف الاقتصادية.

CC : ضعيف للغاية حاليًا وتواجه سنداته أعمل مضاربة شديدة.

C : مفلس تقريبًا ولكنه مستمر في سداد الالتزامات المالية.

يشمل نظام وكالة موديز الأكواد Aaa وAa وA وBaa إلخ، ولكنها متشابه مع التصنيف السابق كثيرًا. فشلت وكالات التصنيف الائتماني فشلاً ذريعًا عند بداية الأزمة المالية العالمية وحيث استمرت في إعطاء بعض التزامات الديون المضمونة (CDOs) أعلى الدرجات بعد فترة قصيرة من الإعلان عن "خطرها الشديد". ولكن نظرًا لعدم تفوق أحد الوكالات على الأآخرتين ونظرًا للدور المهم الذي تلعبه تلك الوكالات في النظام المالي، استمرت في عملها.

10-2 استخدام درجات الائتمان لمخاطر الائتمان

يتم استخدام درجات الائتمان من أجل تقديم توجيهات حول مخاطر تخلف الشركة عن دفع الديون. وسوف تختلف المنهجية الدقيقة المستخدمة: فعند تقييم المخاطر المرتبطة بالسندات الحكومية، ستتخذ الشركة نهجا مختلفا عن ذلك الذي تستخدمه عند تصنيف الشركات، وتتبع نفس الطريقة فيما يتعلق بمخاطر الشركات. وستطلب الشركة التي ترغب في اقتراض المال عن طريق إصدار السندات من إحدى وكالات التصنيف أن يقومون بتصنيفهم، وستقوم بالدفع لهم مقابل هذه الخدمة. في الواقع، فإن أكبر وكالتين وهما ستاندرد آند بورز وموديز، سوف يقومون بتقييم جميع سندات الشركات الكبيرة الصادرة في الولايات المتحدة، سواء طلب ذلك أم لا.

وعلى أية حال يجوز للمصدر أن يدفع الرسوم ويشرع في مشاركة أكثر جدية مع شركة التصنيف لتجنب الوضع الذي يتم فيه التقييم دون معرفة كاملة بالوقائع.

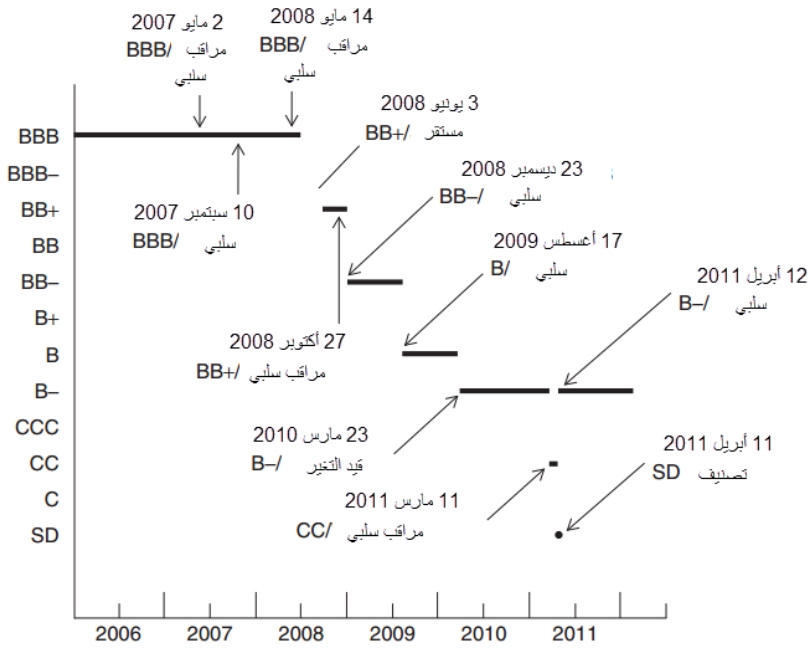
فإن المستثمرين المؤسسيين يفضلون أن يكون إصدار السندات مصنفا من قبل أكثر من شركة، وستنظر شركة التصنيف في أمرين: مخاطر الأعمال والمخاطر المالية.

إن تقييم مخاطر الأعمال ينطوي على النظر في شروط التداول، وتوقعات القطاع، ونوعية فريق الإدارة. ويتم تقييم المخاطر المالية بطريقة أكثر كمية باستخدام البيانات المحاسبية التي تبحث في الربحية ومستويات الدين والقوة المالية ومرونة الشركة. وبعد ذلك سيتم زيارة الشركة من قبل المحللين المعنيين، وستجتمع لجنة للنظر في توصيات المحللين والتصويت على الدرجة النهائية. ويتم إعلام الشركة بالتوصية والتعليق الذي سوف يتم نشره عنها، ومن ثم إعطاء فرصة أخيرة لتقديم معلومات إضافية (أو طلب حذف معلومات سرية للشركة من التعليق) قبل تأكيد التصنيف ونشره.

وإذا كانت الشركة تقوم بإصدار سندات ذات خصائص مختلفة، فيتم تصنيفها بالرجوع إلى التصنيف العام للشركة ومن خلال الديون العليا التي لها أولوية في الدفع، فسيكون من المرجح أن تعطي تصنيفا أعلى من الديون الفرعية. وفي حالة المنتجات عالية التنظيم مثل منتجات CDO يتم دمج الديون من مصادر مختلفة وتقسيمها،ومن ثم تكون العملية أكثر تعقيدا، لأن المستوى العام للمخاطر ينطوي على النظر في الأجزاء المكونة للمنتج المنظم. وبمجرد إعطاء التقييم، تتحول العملية إلى وضع المراقبة حيث يتم تكليف فريق مستقل ضمن وكالة التصنيف بمراقبة التطورات داخل الشركة من أجل أن يكون هناك تغيير في التصنيف في الوقت المناسب (إما أعلى أو أقل) إذا اقتضت الظروف ذلك. وتصدر التصنيفات الفعلية مع علامات زائد أو ناقص ويتم إرفاقها لجميع الدرجات التي تتراوح بين AA وCCC.

وعلاوة على ذلك، تقدم ستاندرد آند بورز ما تصفه ببيان التوقعات الذي يشير إلى اتجاه التغيير إذا كانوا يتوقعون حدوث تغيير خلال العام أو العامين التاليين (بعنوان "التطوير" إذا كان التغيير محتملا ولكن يمكن أن يكون في أي من الاتجاهين). وإذا دخلت الشركة مرحلة حيث تعتقد شركة التصنيف أن هناك فرصة كبيرة لحدوث تغيير في إطار زمني أقصر مدته ثلاثة أشهر أو نحو ذلك، ثم يتم وضع الشركة على حالة "مراقبة الائتمان". ويمكننا أن نرى كيفية تطبيق هذا في مثال ليز كليبورن في الجدول الزمني المبين في الشكل 10-1. ويلاحظ أنه كان هناك تصريحات في مايو 2007، سبتمبر 2007 ومايو 2008 تعكس

إمكانية انخفاض التصنيف بسبب تراكم المشاكل على الشركة ولكن لم يحدث أول انخفاض حتى يونيو 2008.



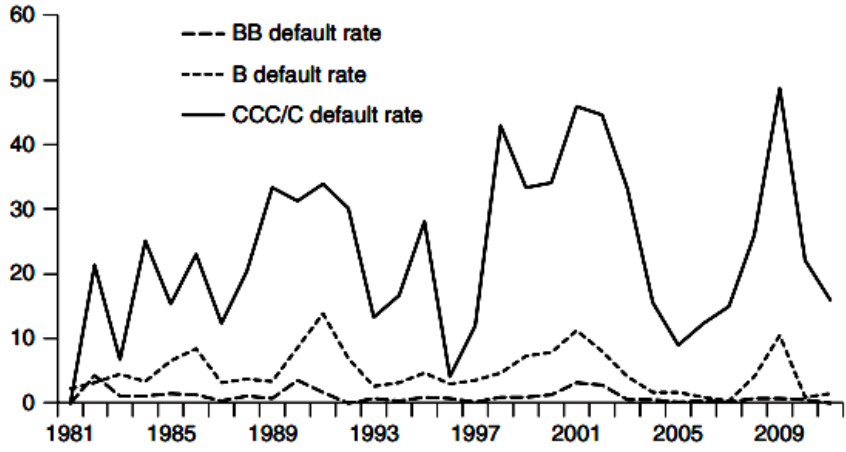
الشكل 10.1 الجدول الزمني لإعلان وكالة ستاندرد آند بورز التصنيفات الائتمانية لشركة *Liz Claiborne Inc*

الشكل 10-1: الجدول الزمني للتصريحات الصادرة عن مؤسسة ستاندرد آند بورز للتصنيف الائتماني لشركة ليز كليبورن.

وتؤدي شركات التصنيف الائتماني مهمة حيوية في سوق رأس المال حيث لديهم خبرة في مهمة تقييم المخاطر وهم مستقلون عن الشركات الفردية. وإن المعلومات الإضافية التي تقدمها للمستثمرين والمشاركين في السوق يكون هدفها في نهاية المطاف أنه يمكن للشركات أن تزيد أرباحها بتكاليف أقل حيث أن التقييمات توفر وسيلة فعالة للمستثمرين للحد من عدم اليقين. وفي الوقت نفسه، فإن شركات التصنيف تقوم بوظيفة مهمة بالنسبة للمنظمين الذين يرغبون في الحد من المخاطر التي يمكن أن تتخذها الشركات المالية ويمكن استخدام التصنيفات ضمن القواعد التي وضعوها. ومن الواضح أن احتمالية التخلف أو

التعسف في السداد تختلف باختلاف عوامل الاقتصاد الكلي: فقد أدت أزمة الائتمان العالمية في الفترة 2008-2009 إلى حدوث الكثير من حالات التخلف عن السداد. ولا يتعين على شركات التصنيف أن تعطي تصنيفات تتماشى مع احتمالات التخلف عن السداد، لأن ذلك سيتطلب إعادة تصنيف كامل مع تغير المناخ الاقتصادي من سنة إلى أخرى. وعوضا عن ذلك، فإن شركات التقييم تهتم بالاحتمالات النسبية: إذا كانت هناك شركة مصنفة بدرجة استثمار فسوف تكون أقل عرضة أن تتخلف عن الدفع بالمقارنة مع شركة مصنفة بدرجة BB والتي بدورها تكون أقل عرضة للتخلف عن شركة مصنفة بدرجة B وهكذا.

ويوضح الشكل 10-2 كيف تتفاوت النسبة المئوية للتخلف عن السداد في السنة مع مرور الوقت بالنسبة للدرجات الثلاثة غير الاستثمارية التي قدمتها شركة ستاندرد آند بورز. وكما تتوقع ستاندرد آند بورز فإن المعدلات الافتراضية لدرجات C هي أعلى من درجات B، والتي بدورها أعلى من درجات BB. ولكن من المثير للاهتمام أيضا أن نرى مدى التباين بينهما. أما بالنسبة للسندات بدرجة C في الأعوام مثل 2005 و1996 ، كان هناك احتمالية بنسبة 10٪ أن يحدث تخلف عن السداد، ولكن في عامي 2001 و2009، إذا تم بدأ العام بدرجة C فإن هذا يعني وجود احتمال بنسبة 40٪ أن يحدث تخلف عن السداد قبل نهاية العام.



الشكل 10-2: نسبة الشركات المتعثرة خلال السنة لثلاث درجات مختلفة (غير استثمارية)

10-2-1 تحليل سلسلة ماركوف للافتراضات.

ولكي نفهم المخاطر المرتبطة بدرجة معينة فإنه يجب أن نحدد الأفق الزمني المعني. هناك فرصة ضئيلة جدا بالنسبة لشركة بدرجة AA أن تتخلف عن السداد هذه السنة.

ولكن إذا تراجع تصنيف الشركة إلى درجة BB في غضون خمس سنوات، فستزيد احتمالية التخلف عن السداد بشكل كبير.

وتقوم مؤسسة ستاندارد & بور بنشر تقريرا يتضمن احتمالات حدوث التغيرات المختلفة الممكنة. ويبين الجدول 10-1 متوسط معدلات التغيرات على مدى 30 عام.

وبالتالي، على سبيل المثال، في المتوسط هناك 8.7٪ من الشركات المصنفة بدرجة AAA تراجعت إلى الدرجة AA خلال العام. ويوضح العمود NR في الجدول النسب الغير مقيمة. وفي بعض الأحيان قد تخرج شركة من التقييمات قبل وقت قصير من التخلف عن السداد، ولكن قد يكون هذا قرار الشركة عندما ترى أنه ليس هناك ضرورة للتصنيف. وستقوم شركة التصنيف بتتبع الشركات التي تخرج من التصنيفات في منتصف العام ثم تتخلف عن السداد خلال ذلك العام، حتى تكون الأرقام المتعلقة بالتخلف السنوي صحيحة.

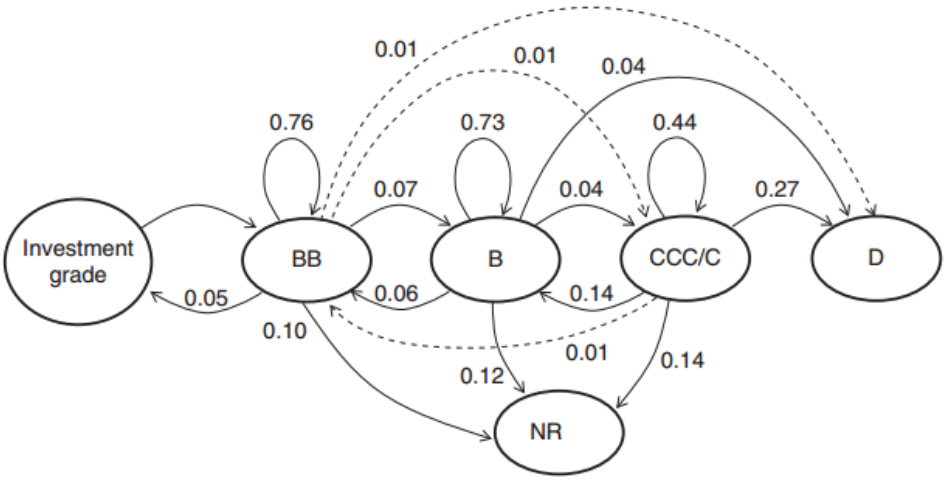
الجدول 10-1: متوسط معدلات التغير في الشركات العالمية للفترة 1981-2011 (%)

	AAA	AA	A	BBB	BB	B	CCC/C	D	NR
AAA	87.2	8.7	0.5	0.1	0.1	0	0.1	0	3.4
AA	0.6	86.3	8.3	0.5	0.1	0.1	0	0	4.1
A	0	1.9	87.3	5.4	0.4	0.2	0	0.1	4.7
BBB	0	0	3.6	84.9	3.9	0.6	0.2	0.2	6.4
BB	0	0	0.2	5.2	75.9	7.2	0.8	0.9	9.8
B	0	0	0.1	0.2	5.6	73.4	4.4	4.5	11.7
CCC/C	0	0	0.2	0.3	0.8	13.7	43.9	26.8	14.4

ومن الصعب عمل رسم بياني يوضح جميع التغيرات الممكنة – وفي الشكل 10-3 أظهرنا التغيرات السنوية بين الدرجات "الغير استثمارية"، وقمنا بتحويل النسب المئوية

إلى احتمالات وتم حذف التغيرات التي لها احتمال أقل من 0.01 أما تلك التي لها احتمال أقل من 0.02 فقمنا بإظهارها بشكل متقطع.

وبالنظر إلى المعلومات المتعلقة بفرص التخلف عن السداد أو إجراء عملية انتقالية في سنة واحدة، فما الذي يعنيه ذلك عن فرصة التخلف عن السداد خلال فترة زمنية مدتها ثلاث أو خمس سنوات؟ إحدى الطرق البسيطة للتعامل مع هذه المشكلة هي تحليل نموذج سلسلة ماركوف لهذه العملية. ويفترض ماركوف أن التغييرات في التصنيف يعتمد فقط على النتيجة الحالية. فإن الشركة التي لديها تصنيف AA لمدة 15 عام، ليست أكثر أو أقل عرضة للانتقال إلى A بالمقارنة مع شركة حققت درجة AA العام الماضي. وأن الشركة التي انخفض تصنيفها من A إلى BBB ومن ثم إلى BB ليست أكثر أو أقل عرضة للانتقال إلى B مرة أخرى بالمقارنة مع شركة تحرك تصنيفها في الاتجاه المعاكس، بعد أن تم تصنيفها بدرجة B العام الماضي ولكنها انتقلت إلى BB.



الشكل 10-3: الاحتمالات السنوية للتغير بين الدرجات غير الاستثمارية. لا تظهر التغيرات مع احتمال أقل من 0.01، وتظهر التغيرات مع احتمال أقل من 0.02 بشكل متقطع.

واحدة من الحقائق الأساسية لسلسلة ماركوف هو أن القوة العاشرة من مصفوفة الانتقال تعطي فرصة إجراء فواصل انتقالية. ولإظهار سبب حدوث ذلك، سنقوم بتحليل

احتمال أن تكون الشركة التي بدأت بتصنيف B هي معرضة للتخلف خلال سنتين باستخدام المعلومات الواردة في الشكل 10-3. (لتبسيط الحسابات نستخدم هذه الأرقام التقريبية بدلا من الأرقام الأكثر دقة الواردة في الجدول 10-1). سنقوم بكتابة p(X, Y) لاحتمال الانتقال من الحالة X إلى الحالة Y. ثم يمكننا أن ننظر في جميع المسارات الممكنة من B إلى D، حيث يمكن أن يحدث انتقال خلال سنة واحدة أو يمكننا الوقوف عند B في السنة الأولى ومن ثم الانتقال إلى D في السنة الثانية.، أو يمكننا الانتقال إلى C في السنة الأولى والانتقال من C إلى D في السنة التالية، وهكذا. فنحصل على:

Pr (من B إلى D خلال عامين)

$$\begin{aligned} &= p(\mathrm{B},\mathrm{D}) + p(\mathrm{B},\mathrm{B})p(\mathrm{B},\mathrm{D}) + p(\mathrm{B},\mathrm{C})p(\mathrm{C},\mathrm{D}) \\ &\quad + p(\mathrm{B},\mathrm{BB})p(\mathrm{BB},\mathrm{D}) \\ &= 0.04 + 0.73 \times 0.04 + 0.04 \times 0.27 + 0.06 \times 0.01 = 0.08. \end{aligned} \tag{1-10}$$

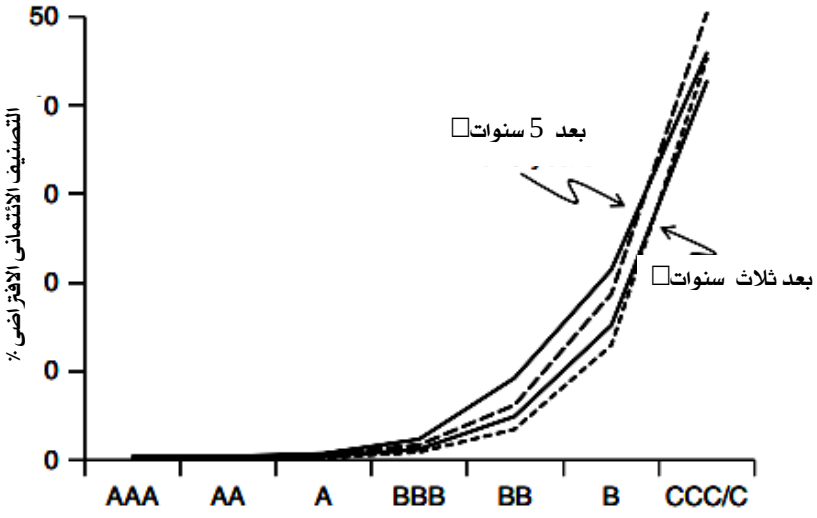
نحن مهتمون باحتمال التخلف عن السداد في وقت ما خلال العامين المقبلين، وبالتالي يمكننا أن ننهي الحسابات بمجرد أن نصل إلى D. مما يجعل من المعقول تحديد $p(\mathrm{D},\mathrm{D}) = 1$ ومن ثم يمكننا أن نرى أن الصيغة لديها النموذج:

$$\Pr(x \text{ to } y \text{ in 2 steps}) = \sum_i p(x,i)p(i,y), \tag{2-10}$$

حيث تشير i إلى كل من الحالات التي يمكننا الحصول عليها من خلال x. في المعادلة (1-10) أخذنا مصطلحات مثل p(B, NR) حيث تكون الاحتمالية في الخطوة الثانية هي صفر. لنفترض أن P تشير إلى المصفوفة المبينة في الجدول 10-1، وتضاف للصفين D وNR حيث في كل حالة هناك احتمال بنسبة 100٪ للبقاء في نفس الحالة. وبالتالي، P هو مصفوفة مربع وسيكون من المناسب لتقسيم كل عنصر بمقدار 100 للتعبير عن احتمالات الانتقال بشكل مباشر بدلا من تمثيله بنسب مئوية. ثم يمكننا كتابة العنصر في الصف i th والعمود j th بالرمز P_{ij}: هو احتمال القفز من الحالة i إلى الحالة j. وتوضح قواعد ضرب المصفوفة أن العنصر في الصف i th والعمود j th لـ $P \times P = P^2$ يعطى بواسطة $\sum_k p_{ik} p_{kj}$ الذي يطابق تماما المعادلة (2-10). وبالتالي، توضح P^2 كل احتمالات الانتقال المكونة من خطوتين،

وتوضح P^3 احتمالات الانتقال المكونة من ثلاثة خطوات وهكذا.

وبما أن شركة ستاندرد آند بورز تقدم تقارير عن التغيرات التي تحدث على مدار ثلاثة سنوات. فيمكننا مقارنة توقعاتنا لمعدلات التخلف عن السداد باستخدام نموذج سلسلة ماركوف للسلوك الفعلي. يتم توضيح مقارنات سلسلة ماركوف الكاملة في جدول البيانات BRMch10-Markov.xlsx ودالة المصفوفة المستخدمة لتنفيذ منتج المصفوفة. ويقارن الشكل 10-4 معدلات التخلف الفعلية والمتوقعة لمدة ثلاث سنوات وخمس سنوات. ويمكننا أن نرى من خلال هذا النموذج أن ماركوف لا يقوم بعمل ممتاز في التنبؤ بهذه المعدلات. ويعد نموذج سلسلة ماركوف وسيلة جيدة للتفكير في حركات التصنيف الائتماني، ولكن يبسط السلوك الحقيقي بشكل كبير. ويمكننا أن نحدد عددا من الأسباب المحتملة للتنبؤات الفقيرة لمدة ثلاث سنوات وخمس سنوات من خلال نموذج ماركوف.



الشكل 10-4: معدلات التخلف الفعلية (الخط الثابت) والمتوقعة (خط متقطع) لمختلف درجات البدء على مدى ثلاث سنوات وخمس سنوات.

1- *تجميع الحالات معاً*. يتم وضع التصنيفات الفعلية بعلامات زائد وناقص، مما يوفر المزيد من الحالات في المجمل. وعندما يتم تجميع الحالات معاً في سلسلة ماركوف، فلا يتم التعامل معها باعتبارها سلسلة ماركوف حيث تتغير خصائصها. على سبيل المثال، إذا تحركت الشركات ببطء من خلال التقييمات على سبيل المثال، +BB إلى BB إلى -BB، إذا كانت الشركة تقع ضمن مجموعة BB لبعض الوقت، فإن هذا قد يزيد من فرصة أن تنتقل إلى المستوى الأدنى -BB، وبالتالي زيادة فرصة القفز إلى المستوى B وكسر افتراض ماركوف.

2- *أنواع مختلفة من الشركات تتصرف بشكل مختلف*. لنفرض أن هناك أنواع مختلفة من الشركات تتبع سلاسل ماركوف المختلفة. فإن هذا لا يعني بالضرورة أن أحد أنواع الشركات هي أكثر خطورة من النوع الآخر (بنفس التصنيف). وربما من خصائص الشركات في القطاع المالي أن تكون أكثر عرضة للتخلف الفوري، بينما يكون الأمر متوازناً بالنسبة للشركات في القطاع غير المالي حيث يوجد خطر أكبر للانتقال إلى الدرجة التالية الأقل. على أية حال، فإن وجود أنواع مختلفة من الشركة يمكن أن يفسد افتراض ماركوف. ويعطي التمرين 10-2مثالا على ذلك.

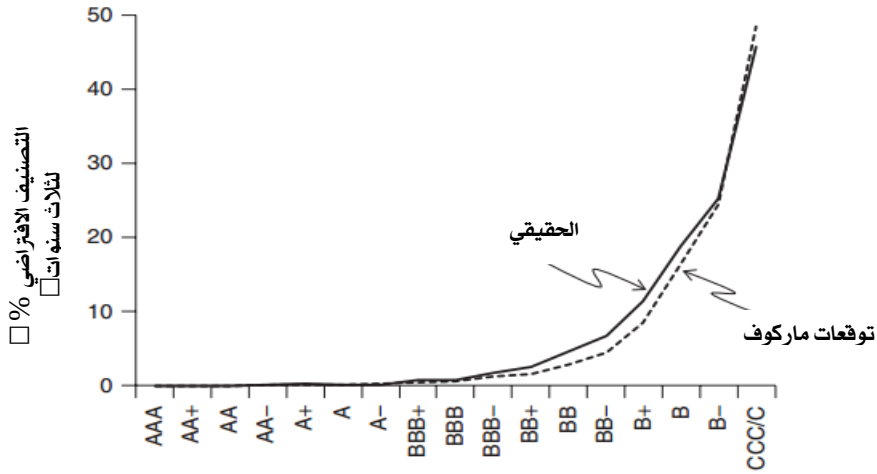
3- *تجميع السنوات السيئة معاً*. وقد لاحظنا بالفعل أن معدلات التخلف عن السداد (والانتقال إلى الدرجات الأقل) تختلف اختلافا كبيرا من سنة إلى أخرى. وهذا يعني أن النموذج الأنسب قد يكون نموذجا يعتمد فيه احتمال حدوث أي انتقال حسب السنة المعنية. ومن ثم يجب أن نستبدل المعادلات كما في المعادلة (10-2)

$$\Pr(x \text{ at time } n \text{ to } y \text{ at time } n+2) = \sum_i p_n(x,i)\,p_{n+1}(i,y)$$

حيث يمثل الجزء الفرعي الفترة الزمنية. إذا كان الاحتمال الكلي لحدوث انخفاض في الدرجات في سنة واحدة يرتبط ارتباطاً إيجابيا باحتمالية التراجع في السنة التالية، فإن ذلك يمكن أن يزيد من الاحتمال الكلي لتراجعين متعاقبين عند مقارنتهما بالحالة عندما تكون الاحتمالات في سنة واحدة مستقلة عن الاحتمالات في السنة التالية.

4- *القرارات الموضوعية التي تتخذها شركة التصنيف*. هناك الكثير من النقاش حول المدى الذي قد تلعب فيه العوامل الذاتية دورا في التصنيف الائتماني. ومن الواضح أن شركات التصنيف نفسها تدعي أن تصنيفاتها موضوعية، ولكن كثيرا ما طرحت أسئلة حول ما إذا كانت التقييمات الأولية سخية جدا، ولأن هناك رغبة كامنة من قبل شركات التصنيف في كسب تجارة مصدري الديون الجديدة إلا أنه قد يكون هناك تردد في إصدار قرار التراجع بسرعة كبيرة، وتنتظر شركة التصنيف حتى تكون على يقين من أنه القرار الصحيح. قد تكون هذه هي الحالة التي تكون فيها أوضاع ديون المصدر مرتبطة بالدرجات. فالتراجع من BBB إلى BB قد يؤدي إلى تسريع سداد الديون، وهذا بدوره قد يؤدي إلى أن تواجه الشركة المزيد من الصعوبات. ومن ناحية أخرى، قد تكون شركات التصنيف متحفظة في تصنيفاتها عمداً، وتحتفظ بالتعديل الذي سيكون ملائما. هذا النوع من السلوك من قبل شركات التصنيف يمكن أن يفسر توقعات مضللة حول افتراض ماركوف.

ويمكننا على الأقل أن نعالج القضيتين الأوليين بشكل أولي باستخدام بيانات ستاندرد آند بورز مقسمة إلى تقييمات أكثر دقة ومقسمة بين الشركات غير المالية والتأمينية والمالية. ويقارن جدول البيانات الثاني في المصنف BRMch10-Markov.xlsx نتائج معدلات التخلف لثلاث سنوات والتي نتجت عن نموذج ماركوف المقتصر على الشركات غير المالية والمعدلات الملاحظة بالفعل، وتظهر النتائج في الشكل 10-5 ويبدو أن تنبؤات ماركوف هي أكثر دقة قليلا، ولكن لا تزال تنطوي على أخطاء كبيرة؛ على سبيل المثال، نموذج ماركوف يتنبأ بمعدل التخلف لمدة ثلاث سنوات بدءا من BB بنسبة 2.8٪ في حين أن الرقم الفعلي الذي لوحظ هو 4.7٪. لاحظ أننا لم ندرج في هذا التحليل أي اعتراف ببيان التوقعات الذي يرافق التصنيف. أحد التفسيرات للمعدلات المرتفعة نسبيا للتخلف عن السداد (التي تم التنبؤ بها من خلال تحليل سلسلة ماركوف) هو أن الشركات التي تم تخفيض تصنيفاتها من المرجح أن تعطى توقعات سلبية (بمعنى آخر، لديهم مخاطر أكبر للتخلف عن السداد من الشركات الأخرى التي لها نفس تصنيف الدرجات).



الشكل 10-5: التوقعات الافتراضية لمدة ثلاث سنوات للشركات غير المالية مقارنة بالأرقام الفعلية المتوسطة.

3-10 الائتمان الاستهلاكي

سوف نركز على الائتمان الاستهلاكي في ما تبقى من هذا الفصل. وهنا يتم تصنيف الفرد بدلا من الشركة. وكان سبب ظهور أول تطبيقات واسعة النطاق للطرق المؤتمتة لتقييم الائتمان هو بسبب طفرة بطاقات الائتمان الذي حدثت في الستينات والسبعينيات. وأدى ذلك إلى قيام المقرضين بالنظر إلى تاريخ الائتمان ومكاتب الائتمان بهدف تجميع البيانات من مختلف الجهات المقرضة حتى يتسنى للمستهلك الذي فشل في سداد بطاقة المتجر، على سبيل المثال، سوف تجد أن المعلومات أصبحت متاحة للمقرضين المحتملين الآخرين (ربما شركة تمويل السيارات).

وتعمل مكاتب الائتمان المختلفة (التي تسمى أحيانا الوكالات المرجعية الائتمانية) في بلدان مختلفة، حيث تقوم مكاتب الائتمان بتجميع معلومات الائتمان من مصادر مختلفة، ويمكنها تقديم تقرير ائتماني عن الفرد عند الطلب، وتصدر الملايين من هذه التقارير كل يوم في جميع أنحاء العالم (أكثر من مليوني يوميا في الولايات المتحدة وحدها). وتعد العملية ببساطة عبارة عن استجواب آلي لقاعدة البيانات. وتختلف البيانات التي تحتفظ بها مكاتب الائتمان

من بلد إلى أخرى وتتأثر بقوانين حماية البيانات. ويتم الاحتفاظ بكمية هائلة من المعلومات في الولايات المتحدة ، في حين أنها تقتصر في بعض البلدان الأوروبية على المعلومات المتاحة للجمهور بشكل رئيسي (على سبيل المثال، سجلات المحكمة للإفلاس). وعادة ما يبقى الدين في السجل فقط إلى أن يتم تسديده، أو حتى انقضاء مهلة محددة (ربما 10 سنوات). وفي العديد من البلدان، يحق للمستهلكين الحصول على نسخة مجانية من سجل الائتمان الخاص بهم.

وتعد أحد التقنيات المستخدمة على نطاق واسع هي دراسة المتقدمين للحصول على القروض ومحاولة الحكم على أساس خصائصهم ومدى احتمال أن يتخلفوا عن السداد. إن مقياس التخلف عن السداد الذي يتم استخدامه عادةً هو احتمال أن يتأخر المتقدم عن السداد لمدة 90 يوماً في غضون الأنثى عشر شهراً المقبلين. ومع ذلك، يعد هذا الإجراء غير مهم لأنه في نهاية المطاف يتم تقييم الأفراد ويتم رفض الأفراد الأقل جاذبية (أي الأكثر احتمالا للتخلف عن السداد) أو تحويلهم إلى نوع مختلف من قروض الائتمان. وتستخدم طريقة مباشرة لتسجيل الائتمان، ويهتم المقرض بما إذا كان سيتم منح الائتمان للفرد، ومن أجل اتخاذ هذا القرار يتم التحقق من عدد من المتغيرات (مثل عدد بطاقات الائتمان التي يتم الاحتفاظ بها، وما إذا كان الفرد مستأجر لمنزله أم مالك) وتستخدم هذه المتغيرات لإعطاء "درجة"، وتستخدم هذه الدرجة أو التقييم للتنبؤ باحتمال تخلف الفرد عن دفع القرض. ويتم أخذ التقييمات من وثيقة التقييم التي تتضمن مكونات التقييم لمختلف الصفات التي يمتلكها الفرد. على سبيل المثال، حقيقة أن مقدم طلب الائتمان قد عاش في نفس العنوان لأكثر من خمس سنوات قد تكون قيمتها 15 نقطة. ويستند تطوير وثيقة التقييم على تاريخ الائتمان لآلاف الأشخاص الآخرين. وإذا كنت مهتما بنتائج التقييم الخاص بك، يمكنك إلقاء نظرة على الموقع: www.myfico.com/ficocreditscoreestimator.

وعادة ما تسمى درجة الائتمان في الولايات المتحدة بـ " تقييم فيكو" (FICO) وفي المملكة المتحدة، هناك موقع أخر على شبكة الإنترنت هو www.checkmyfile.com من خلاله تتمكن من تقييم درجة الائتمان. (والموقع الذي يعادله في أستراليا هو

checkmyfile.com.au). ولم يتم الكشف عن أساليب التقييم بالضبط، ولكن هناك بعض المعلومات المتاحة (على سبيل المثال، 35٪ من نتيجة فيكو ترتبط بتاريخ الدفع). والهدف بالنسبة للمقرض هو ببساطة اتخاذ هذا القرار بأكبر قدر ممكن من الدقة.

وسوف يتم النظر في أي معلومات يمكن استخدامها قانونيا والتي لها تأثير على الجدارة الائتمانية للفرد. ما الذي يعد غير قانوني؟ ولا يجوز التمييز على أساس العرق أو الجنس أو النوع أو الدين، ولذلك لا يمكن طرح هذه الأسئلة. ومع ذلك، فإنه من الجيد النظر في الرمز البريدي للفرد عند تنفيذ الاختيار. وتعد القواعد المتعلقة بالسن والحالة الاجتماعية أكثر تعقيدا وتختلف في البلدان المختلفة. ففي الولايات المتحدة لا يمكن لمصدري البطاقات رفض الائتمان أو تقديم شروط أقل مواتاة على أساس الحالة الاجتماعية.

1-3-10 الاحتمالية، الاحتمالات وتسجيل الاحتمالات

بالنسبة لنوع معين من الأفراد يمكننا استخدام البيانات السابقة للتنبؤ بما إذا كانوا أشخاص جيدة أم لا، أي إذا كانوا سوف يسددون القرض. كتب P_i لهذا الاحتمال لفرد من النوع i. والبديل هو النظر في احتمالات كونها جيدة بدلا من كونها سيئة. هذا قد يكون مألوفا من سياق الرهان: إذا قلنا أن احتمالات فوز الحصان هي 2-1، فإننا نعني أن احتمالية فوز الحصان هي ضعف احتمالية عدم فوزه. الاحتمالات o_i هي ببساطة احتمال كونها جيدة مقسومة على احتمال كونها سيئة، أي

$$o_i = \frac{p_i}{1 - p_i}.$$

احتمالية كونها جيدة تتراوح بين 0 و1. ولكن احتمالات كونها جيدة يمكن أن تختلف من 0 إلى أي قيمة إيجابية. ومن المفيد أيضاً أن ننظر إلى لوجستية الاحتمالات.

$$\log(o_i) = \log\left(\frac{p_i}{1 - p_i}\right).$$

هذه هي اللوغاريتمات الطبيعية التي اتخذت لقاعدة e ويمكن للوغاريتمية الاحتمالات أن تأخذ أي قيمة، سواء إيجابية وسلبية. لاحظ أن كل من الاحتمالات والاحتمالات اللوغاريتمية

يزيدون من دالات الاحتمال. لنفترض أننا نفكر في منح الائتمان إلى فرد. من خلال القيام بذلك سوف نحقق ربحا، ربما من الفائدة التي نتحملها أو ربما لأننا نجني ربحا من منتج لن يتم بيعه إلا إذا قدمنا الائتمان. ومع ذلك، هناك احتمالية حدوث خسارة والتي عادة تكون أكبر بكثير من الربح التي سيحدث إذا كان الفرد لا يسدد الدين. لنفترض أننا نرمز للخسارة بـ L وR للربح، ومن ثم تكون القيمة المتوقعة من القرض بالنسبة لنا هي:

$$p_i R - (1 - p_i)L.$$

هذا يعد أمر إيجابي (بمعنى أننا يجب أن نمضي قدما في القرض) إذا

$$p_i > \frac{L}{L + R},$$

أو ما يعادلها إذا كانت $o_i > (L/R)$، وهو شرط يمكن أن يكون مكتوباً فيما يتعلق بالاحتمالات اللوغاريتمية.

$$\log(o_i) > \log\left(\frac{L}{R}\right).$$

استنادا إلى جميع البيانات التي لدينا، يمكننا تقدير الاحتمالات للفرد المختار بشكل تعسفي. ونقوم بتعريف احتمالات السكان كما يلي:

$$o_{\text{Pop}} = \frac{\Pr(\text{G})}{\Pr(\text{B})}.$$

هذه هي احتمالات أن يكون الفرد جيدا إذا تم اختياره عشوائيا من بين جميع السكان. ونحن نقدر الاحتمالات Pr(G) وPr(B) ببساطة من خلال النظر في النسب الجيدة والسيئة لجميع السكان الذين لدينا بياناتهم.

ونبدأ الآن مناقشة مثال سنواصل استخدامه لتوضيح مناقشتنا في بقية هذا الفصل.

مثال 10-1: بنك سيدني

لنفترض أن بنك سيدني لديه 1200 عميل تم إقراضهم المال على أساس قصير الأجل. وقد استمر معظمهم في تسديد القروض: هؤلاء يسمون "جيدة" أو G، ولم

يستمر بعضهم بالتسديد، ويصنف أي شخص يتخلف عن التسديد لأكثر من 90 يوما بـ "سيء" أو B إلى جانب البيانات المتعلقة بالعمر في وقت اتفاقية القرض، حيث يصنف العمر كما يلي 30، 30–39، 40–49 وأكثر من 50 عام، ويحتفظ بنك سيدني بالبيانات عما إذا كان الأفراد هم ملاك أو مستأجرين، أم لديهم تصنيف آخر فيما يتعلق بمنزلهم، وأيضا ما إذا كان لديهم بطاقة ائتمان أم لا. يتم توضيح البيانات الخاصة بهذا المثال في جدول البيانات BRMch10-BankofSydney.xlsx.

ويمكن تقديم بيانات بنك سيدني بطرق مختلفة. ويبين الجدول 10-2 عدد الأفراد الجيدة والسيئة في كل فئة من الفئات الفرعية التي تم الحصول عليها من "حالة بطاقة الائتمان (2) '×' حالة السكن (3) '×' الفئة العمرية (4) '. في هذه البيانات عدد الأفراد الجيدة هو 1070 وعدد الأفراد السيئة هو 130. وهكذا، بالنسبة لبنك سيدني فإن احتمالات السكان هي $o_{Pop}=1070/130=8.23$. وإذا نظرنا إلى فئة واحدة من الأفراد نجد أن الاحتمالات مختلفة. على سبيل المثال، نظرا لأن الفرد يقل عمره عن 30 عاما، يظهر العمود الأول في بيانات بنك سيدني أن الاحتمالات هي

$$\frac{59+47+63+19+18+12}{5+10+6+2+9+2}=\frac{218}{34}=6.412.$$

لذلك، من المرجح أن يكون المقترضون في هذه الفئة أكثر سوءا من السكان ككل. ويمكننا أيضا إجراء حسابات احتمالات لكل فئة من الفئات الفرعية في الجدول 10-2. وهذا يعطي الأرقام المبينة في الجدول 10-3.

الآن نريد أن نركز على نوع معين من الأفراد، على سبيل المثال، فرد من الفئة A. وتعطى احتمالات الفئة (أ) من خلال:

$$o_A=\frac{\Pr(G\mid A)}{\Pr(B\mid A)}.$$

ويمكن تقدير هذه الاحتمالات من خلال النظر في البيانات، ولكن الآن نركز فقط على الأفراد في الفئة A: نحن ببساطة ننظر في النسب الجيدة والسيئة لهذه المجموعة الفرعية لجميع السكان. وترتبط احتمالات فئة A باحتمالات السكان ككل من خلال شيء يسمى

احتمالات المعلومات 1_A، للفئة A والتي تعرف بأنها:

$$I_A = \frac{\Pr(A \mid G)}{\Pr(A \mid B)}.$$

الجدول 10-2: بيانات عن بنك سيدني: عدد الأفراد الجيدين والسيئين

	Under 30	30–39	40–49	Over 50
مالك بكارت ائتماني	$G = 59$ $B = 5$	$G = 111$ $B = 9$	$G = 118$ $B = 5$	$G = 232$ $B = 11$
مستأجر بكارت ائتماني	$G = 47$ $B = 10$	$G = 16$ $B = 5$	$G = 22$ $B = 4$	$G = 64$ $B = 18$
آخر بكارت ائتماني	$G = 63$ $B = 6$	$G = 21$ $B = 2$	$G = 16$ $B = 3$	$G = 91$ $B = 5$
مالك بدون كارت ائتماني	$G = 19$ $B = 2$	$G = 13$ $B = 3$	$G = 44$ $B = 4$	$G = 31$ $B = 2$
مستأجر بدون كارت ائتماني	$G = 18$ $B = 9$	$G = 14$ $B = 10$	$G = 5$ $B = 1$	$G = 10$ $B = 3$
آخر بدون كارت ائتماني	$G = 12$ $B = 2$	$G = 26$ $B = 8$	$G = 6$ $B = 2$	$G = 12$ $B = 1$

الجدول 10-3: بيانات عن بنك سيدني: الاحتمالات.

	30 تحت	30–39	40–49	50 فوق
مالك بكارت ائتماني	11.80	12.33	23.60	21.09
مستأجر بكارت ائتماني	4.70	3.20	5.50	3.56
آخر بكارت ائتماني	10.50	10.50	5.33	18.20
مالك بدون كارت ائتماني	9.50	4.33	11.00	15.50
مستأجر بدون كارت ائتماني	2.00	1.40	5.00	3.33
آخر بدون كارت ائتماني	6.00	3.25	3.00	12.00

وبناء على مبرهنة باي Pr(G|A)=Pr(A|G)Pr(G)/Pr(A) وبالمثل بالنسبة لـ Pr(B|A) ومن ثم يمكننا إعادة كتابة الاحتمالات للفئة (A) على النحو التالي:

$$o_A = \frac{\Pr(G \mid A)}{\Pr(B \mid A)} = \frac{\Pr(A \mid G)}{\Pr(A \mid B)} \times \frac{\Pr(G)}{\Pr(B)} = I_A \times o_{\text{Pop}}.$$

هنا قمنا بإلغاء المصطلحين Pr(A) للحصول على الصيغة. لذلك، لقد أظهرنا أن تضاعف احتمالات المعلومات لفئة معينة من خلال احتمالات السكان يعطي احتمالات للفئة. توفر احتمالات المعلومات نوع من التعديل لاحتمالات السكان للوصول إلى احتمالات في فئة معينة. إن ترجيح الأدلة (WoE) لفئة هو مجرد اللوغاريتم الطبيعي لاحتمالات المعلومات لهذه الفئة:

$$w_A = \log(I_A) = \log\left(\frac{\Pr(A \mid \mathrm{G})}{\Pr(A \mid \mathrm{B})}\right),$$

ومن الطبيعي أن نستخدم هذا التعريف لأنه يتيح لنا العثور على الاحتمالات اللوغاريتمية لفئة معينة ببساطة عن طريق إضافة ترجيحات الأدلة والاحتمالات اللوغاريتمية للسكان.

$$\begin{aligned}\log(o_A) &= \log(I_A \times o_{\mathrm{Pop}}) = \log(I_A) + \log(o_{\mathrm{Pop}}) \\ &= w_A + \log(o_{\mathrm{Pop}}).\end{aligned}$$

مثال 10-1 (متابعة) بنك سيدني

وسوف نستمر في النظر إلى الأفراد الذين تقل أعمارهم عن 30 عاما، وهو ما يقابل العمود الأول في الجدول 10-2. وكما رأينا من قبل، فإن احتمالات هذه الفئة هي 218/ 34=6.412. ويمكننا أيضا حساب احتمالات المعلومات لهذه الفئة.

وتجدر الإشارة إلى أنه من بين 1070 شخصا جيداً، يوجد 218 شخص في هذه الشريحة العمرية، ومن أصل 130 شخصا سيئا، هناك 34 في هذه الفئة العمرية. بعبارة أخرى:

$$\begin{aligned}\Pr(\text{age} < 30 \mid \mathrm{G}) &= 218/1070, \\ \Pr(\text{age} < 30 \mid \mathrm{B}) &= 34/130.\end{aligned}$$

وبالتالي، فإن احتمالات المعلومات للفئة العمرية أقل من 30 هي

$$I_{\text{age}<30} = \frac{218/1070}{34/130} = \frac{218}{34} \times \frac{130}{1070} = 0.779.$$

و يمكننا حساب ترجيحات الأدلة ل "سن أقل من 30" كما يلي

$$w_{\text{age}<30} = \log\left(\frac{218}{34} \times \frac{130}{1070}\right) = \log(0.779) = -0.2497.$$

وتعني القواعد اللوغاريتمية أن الاحتمالات اللوغاريتمية لهذه الفئة هي

$$\begin{aligned}\log\left(\frac{218}{34}\right) &= \log\left(\frac{218}{34}\times\frac{130}{1070}\right)+\log\left(\frac{1070}{130}\right)\\ &= w_{\text{age}<30}+\log(o_{\text{Pop}})\\ &= -0.2497+2.1079=1.8582.\end{aligned}$$

ولقد أجرينا التجربة السابقة على أفراد مطابقة تماما لهذه الشروط، فمن المنطقي محاولة إجراء اقتطاعات من ما لاحظناه في الفئات الأكبر التي ينتمي إليها هذا الشخص. اتضح أنه إذا كان السلوك مستقل تحت سمات مختلفة، فيمكن جمع ترجيحات الأدلة معاً لإيجاد الاحتمالات اللوغاريتمية للفرد الذي لديه عدد من الصفات المختلفة، ولمعرفة سبب صحة ذلك، ضع في اعتبارك فردا له سمتان A_1 وA_2 نود أن نحسب احتمالات أن يكون هذا الفرد جيد، ونريد إيجاد $\Pr(\text{G} \mid A_1, A_2)$/ $\Pr(\text{B} \mid A_1, A_2)$ ونحن نعلم من خلال قاعدة بايز أنه،

$$\Pr(\text{G} \mid A_1, A_2) = \Pr(A_1, A_2 \mid \text{G})\frac{\Pr(\text{G})}{\Pr(A_1, A_2)}.$$

إذا كان A_1 وA_2 مستقلين (بحيث لا تخبرنا معلومات سمة واحدة أي شيء عن السمة الأخرى)، عندئذ يكون لدينا $\Pr(A\ 1,A2\ |B)=Pr(A1|B)\times\ Pr(A2|B)$ و$\Pr(A1,A2|G)=Pr(A1|G)\times Pr(A2|G)$ ومن ثم تكون احتمالات الفرد في الفئة A_1 وA_2 هي

$$\begin{aligned}\frac{\Pr(\text{G} \mid A_1, A_2)}{\Pr(\text{B} \mid A_1, A_2)} &= \frac{\Pr(A_1, A_2 \mid \text{G})}{\Pr(A_1, A_2 \mid \text{B})}\frac{\Pr(\text{G})}{\Pr(\text{B})}\\ &= \frac{\Pr(A_1 \mid \text{G})}{\Pr(A_1 \mid \text{B})}\times\frac{\Pr(A_2 \mid \text{G})}{\Pr(A_2 \mid \text{B})}\times\frac{\Pr(\text{G})}{\Pr(\text{B})}.\end{aligned}$$

ويمكن توسيع نطاق نفس التعبير إلى أي عدد من المصطلحات. وبذلك تكون الاحتمالات اللوغاريتمية لـ $A_1,A_2, \ldots, A_n$ هي

$$\log\left(\frac{\Pr(\text{G} \mid A_1, A_2, \ldots, A_n)}{\Pr(\text{B} \mid A_1, A_2, \ldots, A_n)}\right) = w_1 + w_2 + \ldots + w_n + \log(o_{\text{Pop}})$$

حيث تمثل o_{Pop} احتمالات السكان وتمثل w_i ترجيحات الأدلة للسمة A_i.

وهذا يقودنا إلى فكرة وثيقة التقييم إذا كنا نعرف ترجيحات الأدلة لكل فئة رئيسة، فبالنسبة للفرد الذي يقع عند تقاطع عدد من فئات مختلفة، عندئذ يمكننا إضافة أرقام w_j ذات الصلة. وأيضاً إضافة مصطلح ثابت معطى من قبل لوغاريتم احتمالات السكان للحصول على التنبؤ للاحتمالات اللوغاريتمية. وهذا ما يسمى وثيقة تقييم بايز وبشكل عام، فإن وثيقة التقييم لديها بنية بسيطة جدا: أنها تربط كل فئة بدرجة معينة ومن ثم تجمع درجات الفرد معاً للحصول على النتيجة النهائية. الآن نعود إلى مثال بنكك سيدني.

مثال 10-1 (متابعة) بنك سيدني

لقد رأينا بالفعل أن $w_{age<30}=-0.2497$. ويمكننا حساب ترجيحات الأدلة الأخرى بنفس الطريقة. على سبيل المثال، هناك 196 شخص جيد بين المستأجرين، و60 شخص سيئ، لذلك:

$$w_{\text{renter}} = \log\left(\frac{196}{60} \times \frac{130}{1070}\right) = \log(0.3969) = -0.9241.$$

لدينا أيضا مصطلح ثابت من $\log(o_{\text{Pop}}) = 2.1079$. وتعطينا هذه القيم w وثيقة التقييم ولكن من المناسب أن تقوم بضرب كل الأرقام في 100 وأخذ أقرب عدد صحيح. وهذا يعطينا وثيقة التقييم الواردة في الجدول 10-4. باستخدام هذا نرى أنه، على سبيل المثال، درجة للفرد الذي يملك منزلاً وبطاقة الائتمان وعمره 42 هي 30+62+23+211=326، في حين أن درجة الشخص البالغ من العمر 28 عاما والذي يعيش في منزل وليس لديه بطاقة الائتمان هي −25+3−61+211=128. ويمكننا التحقق من مدى فعالية هذا باستخدام وثيقة التقييم للتنبؤ باحتمالات مختلف فئات الأفراد. وترد الاحتمالات اللوغاريتمية الفعلية للبيانات في الجدول 10-5.

في مثال صاحب المنزل البالغ من العمر 42 عاما ولديه بطاقة ائتمان، كانت درجته في وثيقة التقييم هي 326، لذلك الاحتمالات اللوغاريتمية المتوقعة هي 3.26، ومن ثم، التنبؤ لاحتمالات هذا الفرد بأن تكون جيدة هي:

$$e^{3.26} = 26.05.$$

الجدول 10-4: وثيقة التقييم المستمدة من ترجيح الأدلة.

السمات	الدرجة
السن <30	−25
السن 30−39	−42
السن 40−49	30
السن ≥50	29
يملك منزل	62
يستأجر منزل	−92
آخر	3
لديه بطاقة ائتمان	23
ليس لديه بطاقة ائتمان	−61
ثابت	211

الجدول 10-5: بيانات بنك سيدني: الاحتمالات اللوغاريتمية.

	30 تحت	30−39	40−49	50 فوق
مالك ببطاقة ائتمان	2.47	2.51	3.16	3.05
مستأجر ببطاقة ائتمان	1.55	1.16	1.70	1.27
آخر ببطاقة ائتمان	2.35	2.35	1.67	2.90
مالك بدون بطاقة ائتمان	2.25	1.47	2.40	2.74
مستأجر بدون بطاقة ائتمان	0.69	0.34	1.61	1.20
آخر بدون بطاقة ائتمان	1.79	1.18	1.10	2.48

في هذه الفئة هناك 118 شخصا جيدا و5 أشخاص سيئة، وبالتالي فإن الاحتمالات الفعلية هي 23.6 إلى 1 وبنفس الطريقة، يمكننا أن ننظر إلى احتمالات الشخص البالغ من العمر 28 عاما ويعيش في منزل وليس لديه بطاقة ائتمان بأنها "جيدة". وتكون النتيجة هي 128، حيث تتفق مع الاحتمالات اللوغاريتمية 1.28، مما يعني أنه من المتوقع أن تكون الاحتمالات:

$$e^{1.28} = 3.60.$$

والنتائج الفعلية في هذه الفئة هي 12 فردا جيدين و2 من الأفراد السيئين، مع وجود احتمالات من 6 إلى 1. وبالتالي، فإن التنبؤ في الحالة الأولى معقول، ولكن في الحالة الثانية

يكون التنبؤ أقل إرضاءً، ومع ذلك، حتى في الحالة الثانية فإن الأمور ليست سيئة للغاية، لأنه إذا كان هناك فرد واحد فقط من الأفراد الحاليين مصنف على أنه سيء، فإن هذا من شأنه أن يحول الاحتمالات إلى 4 إلى 1، وهو قريب من التقييم 3.60.

4-10 الانحدار اللوجستي

وهناك طريقة أفضل للتنبؤ باحتمالات الأفراد من خلال استخدام الانحدار اللوجستي، والذي هو موضوع هذا القسم. وتعد هذه هي الطريقة الأكثر شيوعا لعمل وثيقة التقييم. ولكننا نبدأ بالمسألة الأساسية المتعلقة بكيفية التنبؤ بالاحتمالات. ويتوقع الانحدار العادي متغير تابع y على أساس مجموعة من المتغيرات التفسيرية $j = 2, \ldots, m$ ويمكن كتابة الشكل الخطي لهذا التنبؤ هكذا:

$$y = \beta_0 + \sum \beta_i x_i$$
$$= \beta \cdot \mathbf{x}$$

حيث نستخدم حرف بارز للإشارة إلى المتجه: $x=(x_0,x_1, \ldots, x_m)$, $\beta=(\beta_0,\beta_1, \ldots, \beta_m)$ ونقيم $x_0 = 1$ ، ثم نقيم βs من خلال النظر في البيانات السابقة، ولا يمكننا استخدام نفس النهج عند التعامل مع الاحتمالات، لأننا نريد أن نضمن أن التنبؤات y كلها بين 0 و1.

النهج اللوجستي هو استخدام تحول غير خطي للحصول على التنبؤ الاحتمالي p من قيم x_i:

$$p = \frac{e^{\beta \cdot \mathbf{x}}}{1 + e^{\beta \cdot \mathbf{x}}}. \qquad (3\text{-}10)$$

إذا أصبحت قيمة β·x كبيرة جداً، إذن فالنهوج P تساوي 1، إذا كانت β·x=0، إذن p=1/2 ، وإذا كانت β·x هي رقم سلبي كبير، إذن فالنهوج 0 ، ويبين الشكل 10-6 ما تبدو عليه هذه الدالة (وهناك وظائف أخرى قد نختارها والتي من شأنها أن تنتج نتائج مماثلة). وميزة الدالة اللوجستية هي أنه تناسب تعريفنا السابق للاحتمالات اللوغاريتمية بشكل جيد. حيث لدينا:

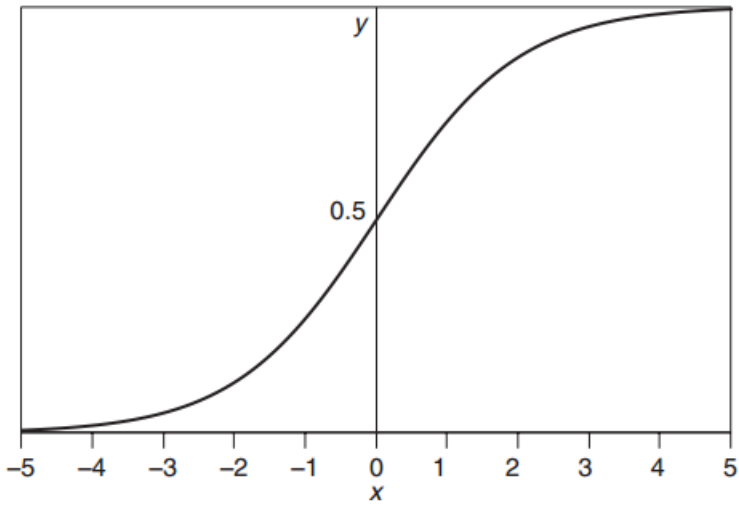
$$1 - p = 1 - \frac{e^{\beta \cdot \mathbf{x}}}{1 + e^{\beta \cdot \mathbf{x}}} = \frac{1}{1 + e^{\beta \cdot \mathbf{x}}}.$$

إذن

$$\text{odds} = \frac{p}{1-p} = \frac{e^{\beta \cdot \mathbf{x}}}{1+e^{\beta \cdot \mathbf{x}}} \frac{1+e^{\beta \cdot \mathbf{x}}}{1} = e^{\beta \cdot \mathbf{x}},$$

وبالتالي

$$\log_e \left(\frac{p}{1-p} \right) = \log_e(\text{odds}) = \log_e(e^{\beta \cdot \mathbf{x}}) = \beta \cdot \mathbf{x}. \quad (4\text{-}10)$$



الشكل 10-6: رسم بياني للدالة اللوجستية $y = e^x/(1+e^x)$

وبالتالي، إذا كانت الاحتمالات الفعلية للجيد أو السيئ تعطى بواسطة نموذج لوجستي مستمد من دالة خطية $\beta \cdot x$، فإن الاحتمالات اللوغاريتمية سوف تكون دالة خطية للملاحظات x_i

كيف يمكننا تقدير القيم β_0، β_1، ...، β_m؟ ففي الانحدار العادي نحاول تقليل مصطلح الخطأ من خلال النظر في التنبؤ ضد ما حدث فعلا. ولكن بالنسبة لنموذجنا اللوجستي، فإن التنبؤ هو احتمال وما حدث فعلا قد يكون نتيجة جيدة أو نتيجة سيئة (والتي يمكن تكون إما 1 أو 0). هناك خياران: نهج بسيط باستخدام انحدار المربعات الصغرى العادية

أو أسلوب الانحدار اللوجستي الأكثر تطورا. وسنتعامل مع النهج الأبسط أولا.

يعمل انحدار المربعات الصغرى مع المعادلة (10-4) بشكل مباشر. ونظرا لأن الاحتمالات اللوغاريتمية هي مزيج خطي من المتغيرات التفسيرية، فيمكننا تقييم المعاملات مباشرة باستخدام انحدار المربعات الصغرى العادية استنادا إلى الاحتمالات اللوغاريتمية لكل فئة. وتعمل هذه الطريقة تعمل بشكل جيد في الممارسة العملية، وميزتها أنه يمكن استخدامها دون استخدام أداة تحليل البيانات التي تتضمن عنصر الانحدار اللوجستي (على سبيل المثال، يمكننا استخدام انحدار المربعات الصغرى من خلال تطبيق "تحليل البيانات" الإضافية التي توجد في إكسيل). نحن بحاجة إلى توفير المزيد من المعلومات حول صيغة المتغيرات التفسيرية.

ومن الناحية العملية، فإن أهم حالة هي عندما تكون المتغيرات التوضيحية فئوية. وفي بعض الأحيان ينتج هذا من طبيعة البيانات (على سبيل المثال، هل الفرد مستأجر لمنزله أم لا؟). ولكن حتى عندما يكون هذا غير صحيح، فإن معظم درجات الائتمان تعطي البيانات الفئوية من خلال تقسيم الأفراد إلى فئات. والمثال الأكثر وضوحا هو متغير مثل العمر: وهذا يعامل بشكل طبيعي كمتغير مستمر، ولكن تحديد النقاط الائتمانية عادة ما يحدد بعض الفئات العمرية وتخصيص الأفراد إلى واحد فقط من هذه الفئات. وينطبق نفس الشيء على مستويات الدخل. ففي هذه الحالة نستخدم متغيرات وهمية لكل فئة.

لذلك، على سبيل المثال، إذا كنا نستخدم الأقواس العمرية: (A) أقل من 30، (B) من 30 إلى 39، (C) 40 إلى 49، (D) 50 أو أكثر، فيكون الشخص البالغ من العمر 33 عاماً لديه $x_{ageA}=1-x_{ageB}-x_{ageC}-x_{ageD}$ حيث أن $x_{ageA}=0, x_{ageB}=1, x_{ageC}=0, x_{ageD}=0$ ، وتؤدي هذه المتغيرات الفئوية تلقائيا إلى مشكلة الخطية المتداخلة في الانحدار. وبعبارة أخرى، لا توجد معلومات إضافية في الفئة الأخيرة.

ويعد أفضل نهج هو تنفيذ الانحدار دون ظهور واحدة من الفئات. لذلك، إذا تركنا x_{ageA} في هذا المخطط، فسوف يتم تسجيل الفرد البالغ 25 عام كالتالي $x_{ageB}=x_{ageC}=x_{ageD}=0$. وبالطبع هذا هو الشيء الطبيعي الذي يجب القيام به عندما يكون للمتغير

خيارين فقط. ومع مثل هذا التصنيف من المتغيرات التفسيرية من الطبيعي أن يكون هناك عدد من الأفراد في كل من الخلايا الممكنة أي لكل مجموعة ممكنة من الفئات. فإن مجموعة الأفراد في خانة معينة والذين يكون بعضهم جيد وبعضهم سيئ، تمكننا من حساب قيمة الاحتمالات اللوغاريتمية لمجموعة من المتغيرات الوهمية المقابلة لتلك الخانة. وتصبح قيم الاحتمالات اللوغاريتمية هي المتغير التابع في تحليل الانحدار.

و هناك نقطتين ينبغي أخذهما في الحسبان عند استخدام هذا النهج:

- تعطي هذه الطريقة احتمال متساو لجميع الفئات بغض النظر عن عدد الأفراد المعنيين. في الواقع، يجب أن نكون أكثر قلقا حول الأخطاء في تنبؤات الاحتمالات اللوغاريتمية عندما يكون هناك عدد كبير من الأفراد في فئة واحدة.
- سوف يفشل نهج الانحدار المربعات الصغرى إذا لم يتم تعريف الاحتمالات اللوغاريتمية لفئة محددة حيث لن يوجد أفراد "سيئون" في تلك الفئة. في حالة ما إذا أصبحت الاحتمالات اللوغاريتمية لانهائية فإن الإجراء الموصى به هو إضافة 0.5 إلى كل من الأرقام الجيدة والسيئة للحصول على قيمة الاحتمالات اللوغاريتمية المعدلة. لذلك، على سبيل المثال، إذا كان لدى مجموعة معينة 10 أفراد فقط، كل منهم "جيد"، قد نستبدل قيمة الاحتمالات اللوغاريتمية اللانهائية بـ $\log(10.5/0.5)=3.0445$ (لاحظ أن هذا لا يزال أعلى بكثير من قيمة الاحتمالات اللوغاريتمية إذا كان 9 من أصل 10 أفراد "جيدون"، ويكون $\log(9/1)=2.1972$).

الآن ننتقل إلى نهج أكثر تطورا (موصي به) من الانحدار اللوجستي. والفكرة هنا هي استخدام تقييم أقصى احتمالية للمعاملات β. وبعبارة أخرى، فإننا ننظر إلى بيانات النتائج الفردية الجيدة والسيئة ونفترض أن احتمال الحصول على النتائج الجيدة سيكون من خلال المعادلة (10-3). ومن ثم نريد أن نختار قيم β التي تزيد من احتمال مجموعة الملاحظات الخاصة بنا (جميعها إما 0 أو 1). ولنرى كيف يتم عمل ذلك، نفترض أن هناك فردين نريد تقديم ملاحظتين لهما y_1 وy_2. لنفترض أن لمجموعة معينة من القيم β نحسب احتمال أن $y_1=1$ (جيد) وتحويلها إلى p_1، وبنفس الطريقة، نحسب احتمال أن $y_2=1$

وتحويلها إلى p_2. ومن هنا يمكننا حساب احتمالية ملاحظة قيم y_1 وy_2 التي لاحظناها. وهذا ما يسمى الاحتمال المرتبط بقيم β. على سبيل المثال، يتم الحصول على احتمالية ملاحظة $y_1 = 1$ و$y_2 = 0$ من خلال $p_1(1-p_2)$. إذا كانت القيم x المرتبطة بالملاحظة الأولى هي المتجه $x^{(1)}$، وللمراقبة الثانية المتجه $x^{(2)}$، لدينا:

$$p_1(1-p_2) = \frac{e^{\beta \cdot \mathbf{x}^{(1)}}}{1+e^{\beta \cdot \mathbf{x}^{(1)}}} \left(1 - \frac{e^{\beta \cdot \mathbf{x}^{(2)}}}{1+e^{\beta \cdot \mathbf{x}^{(2)}}}\right). \quad (5\text{-}10)$$

وتتمثل المرحلة التالية في تقدير أقصى احتمال للمتجه β عن طريق اختيار القيمة التي تزيد من هذا التعبير. والفكرة هي أن الملاحظات التي أجريناها تجعل بعض القيم β غير محتملة.

وبشكل عام عادةً ما يكون لدينا العديد من الملاحظات n وليس فقط أثنين، ولكن المبدأ هو نفسه: سيكون هناك تعبير طويل لمنتج ينطوي على قيمة واحدة β وقيم مختلفة لمتغيرات الفئة x. وينطوي الانحدار اللوجستي على تعظيم هذا التعبير لجميع الخيارات الممكنة لـ β. وسنقدم المزيد من التفاصيل حول هذا الموضوع في القسم التالي. ولكن في كل الحالات يتطلب الانحدار اللوجستي استخدام البرمجيات لأغراض خاصة. وقد تم الحصول على الانحدار اللوجستي لمثال بنك سيدني أدناه باستخدام البرمجيات الحرة "غريتل".

إن جمال الانحدار اللوجستي هو أنه يعمل بشكل جيد عندما يكون هناك العديد من أنواع السمات والخصائص بدلاً من وجود ثلاث سمات فقط كما في مثال بنك سيدني (بطاقة الائتمان، نوع السكن، شريحة العمر). إذا قمنا بقياس الأفراد من خلال إجاباتهم، على سبيل المثال، 10 أسئلة مختلفة، فحتى إذا كانت الأسئلة بسيطة (نعم أو لا) سوف ينتهي بنا المطاف بتجميع أكثر من 1000 خانة. وهذا يعني أنه سيكون هناك العديد من الخانات التي لا يوجد بها أفراد وأيضاً العديد من الخانات حيث لا يوجد أفراد "سيئة". ولكن ما زلنا بحاجة إلى طريقة لتقييم مخاطر الائتمان لعميل جديد قد ينتمي إلى واحدة من الخانات حيث أننا ليس لدينا العديد من العملاء السابقين للمقارنة بينهم.

مثال 10-1 (متابعة) بنك سيدني

بالنسبة لمثال بنك سيدني، فإن الانحدار اللوجستي يأخذ في الاعتبار الملاحظات الفردية التي يقدر عددها 1200 وتقييم قيم β المرتبطة بالمتغيرات الستة: x_{own}= "المالك"، x_{rent} = "المستأجر"، x_{cc} = "حامل بطاقة الائتمان"، x_{ageB}= "العمر من 30 إلى 39"، x_{ageC} = " من 40 إلى 49"، x_{ageD} = "أكثر من 50". أما المتغيرات الثلاثة الأخرى هي: "أخرى"، "بدون بطاقة ائتمان"، "العمر تحت 30" وتعتبر هذه المتغيرات مجموعات خطية من المتغيرات التفسيرية الأخرى ولا تضيف أي شيء إلى الانحدار، لذلك يتم استبعادها. وهكذا، على سبيل المثال، فإن الأفراد في أعلى الخانة اليسرى من الجدول 10-2 لديهم x_{own}=1، x_{rent} =0، x_{cc} = 1، x_{ageB} =0، x_{ageC}=0 ، x_{ageD} =0. وهكذا، فإن المنتج العددي لكل من هؤلاء الأفراد $\beta \cdot x = \beta_{own} + \beta_{cc}$ والتي تمثل أيضاً الاحتمالات اللوغاريتمية. الاحتمال المتوقع لكون الـ 64 شخصا بأن يكونوا "جيدين" هو:

$$p = \frac{e^{\beta_{own}+\beta_{cc}}}{1 + e^{\beta_{own}+\beta_{cc}}}.$$

هذه هي القيمة التي نحاول أن تتطابق مع القيم الفعلية من خلال الاختيار الصحيح للقيم β، تم استخدام البيانات الواردة في جدول البيانات BRMch10-Bank of Sydney.xlsx (الورقة الثانية) لتفعيل الانحدار اللوجستي (باستخدام البرمجيات الحرة غريتل). ويظهر جزء من المخرجات في الجدول 10-6. وهكذا، لدينا قيم المفردات على النحو التالي:

$$\begin{aligned}
\beta_{cc} &= 0.629 \\
\beta_{own} &= 0.493 \\
\beta_{rent} &= -0.984 \\
\beta_{ageB} &= -0.370 \\
\beta_{ageC} &= 0.134 \\
\beta_{ageD} &= 0.204 \\
\beta_0 &= 1.723
\end{aligned}$$

حيث β_0 هو الدالة الثابتة. وبالتالي، نحصل على تقييم أقصى احتمالية:

$$\log_e\left(\frac{p}{1-p}\right) = 1.723 + 0.629x_{cc} + 0.493x_{own} - 0.984x_{rent} - 0.370x_{ageB} + 0.134x_{ageC} + 0.204x_{ageD}. \quad (6\text{-}10)$$

الجدول 10-6: الناتج من الانحدار اللوجستي

	الدرجة	الخطأ القياسي
ثابت	1.72256	0.287793
بطاقة ائتمان	0.628808	0.211980
مالك	0.493278	0.261849
مستأجر	−0.983520	0.249570
السن من 30 إلى 39	−0.369558	0.271993
السن من 40 إلى 49	0.134251	0.319367
السن من 50 فيما فوق	0.204348	0.259001

يمكننا استخدام هذه القيم لإنتاج جدول الاحتمالات اللوغاريتمية لكل من الفئات الـ 24. ويظهر ذلك في الجدول 10-7 ، ويمكن مقارنته بالبيانات الواردة في الجدول 10-5. يمكننا أيضا مقارنة هذه النتائج مع ما يحدث إذا استخدمنا انحدار المربعات الصغرى، إذا فعلنا ذلك مع بيانات بنك سيدني، نحصل علي:

$$\log_e\left(\frac{p}{1-p}\right) = 1.650 + 0.575x_{cc} + 0.527x_{own} - 0.788x_{rent} - 0.349x_{ageB} + 0.090x_{ageC} + 0.424x_{ageD}.$$

الجدول 10-7: بيانات بنك سيدني: التنبؤ بالاحتمالات اللوغاريتمية من الانحدار اللوجستي.

	30 تحت	30–39	40–49	50 فوق
مالك ببطاقة ائتمان	2.84	2.48	2.98	3.05
مستأجر ببطاقة ائتمان	1.37	1.00	1.50	1.57
آخر ببطاقة ائتمان	2.35	1.98	2.49	2.56
مالك بدون بطاقة ائتمان	2.22	1.85	2.35	2.42
مستأجر بدون بطاقة ائتمان	0.74	0.37	0.87	0.94
آخر بدون بطاقة ائتمان	1.72	1.35	1.86	1.93

وبالنسبة للفئات الفرعية التي يوجد فيها عدد كبير من الأفراد، فإننا نتوقع أن نرى تنبؤات الاحتمالات اللوغاريتمية من الانحدار اللوجستي أفضل من استخدام نهج المربعات الصغرى العادية. وينطبق ذلك على بيانات بنك سيدني، حيث أن نتيجة تنبؤات الاحتمالات اللوغاريتمية من الانحدار اللوجستي هي 3.05 للنوع الأكثر شيوعا من العملاء (فوق الـ 50 عام ومالك المنزل وحامل بطاقة الائتمان)، في حين يعطي انحدار المربعات الصغرى البديلة تنبؤات الاحتمالات اللوغاريتمية:

$$1.650 + 0.575 + 0.527 + 0.424 = 3.18$$

وهي مرتفعة قليلا.

* 10-4-1 مزيد من التفاصيل عن الانحدار اللوجستي

لقد رأينا في الملاحظتين Y_1 وY_2 ومع احتمال أن $y_i = 1$ المعطى من P_i، واحتمال أن $y_1 = 1$ و$y_2 = 0$ هو $p_1(1-p_2)$، يمكننا تعميم هذا التعبير على أي مجموعة من القيم لـ y_1 وy_2 من خلال الإشارة إلى أن الاحتمال يعطى من قبل:

$$(p_1)^{y_1}(1-p_1)^{(1-y_1)}(p_2)^{y_2}(1-p_2)^{(1-y_2)}.$$

وبشكل أعم، يمكننا أن نأخذ الملاحظات N، y_i ,i =1,2, ...,n (جميع y_i قيمتها 1 أو 0).

$$\prod_{i=1}^{n}(p_i)^{y_i}(1-p_i)^{(1-y_i)}. \qquad (7\text{-}10)$$

هذا هو احتمال الملاحظة 1، y_2، ...y_n، إذا كانت الاحتمالات $p_1,p_2,...p_n$. إذا كانت p_i تأتي من نموذج لوجستي، فإن:

$$p_i = \frac{e^{\beta \cdot \mathbf{x}^{(i)}}}{1 + e^{\beta \cdot \mathbf{x}^{(i)}}}, \qquad (8\text{-}10)$$

حيث قمنا بكتابة $\mathbf{x}^{(i)} = (x_0^i, x_1^i, \ldots, x_m^i)$ للمتغيرات المستقلة المرتبطة بهذه الملاحظة. لذا فإن مشكلة إيجاد تقدير أقصى احتمالية من مجموعة معينة من البيانات (y_1,y_2, ...y_n) يعادل إيجاد مجموعة من القيم β التي تعظم تعبير المنتج (7-10) حيث تعطى p_i بالمعادلة

(8-10). ولكن إذا أردنا تعظيم تعبير، فيمكننا أيضا تعظيم لوغاريتم هذا التعبير. لذلك، بدلا من تعظيم الاحتمال، نقوم بتعظيم احتمال اللوغاريتم.

$$L = \log\left(\prod_{i=1}^{n} (p_i)^{y_i}(1-p_i)^{(1-y_i)}\right)$$
$$= \sum_{i=1}^{n} y_i \log(p_i) + \sum_{i=1}^{n}(1-y_i)\log(1-p_i).$$

وبما أن قيم P_i تعطى بدالات لوجستية، فيمكن تبسيطها إلى حد كبير:

$$L = \sum_{i=1}^{n} y_i \log\left(\frac{e^{\beta\cdot\mathbf{x}^i}}{1+e^{\beta\cdot\mathbf{x}^i}}\right) + \sum_{i=1}^{n}(1-y_i)\log\left(\frac{1}{1+e^{\beta\cdot\mathbf{x}^i}}\right)$$
$$= \sum_{i=1}^{n} y_i(\log(e^{\beta\cdot\mathbf{x}^i}) - \log(1+e^{\beta\cdot\mathbf{x}^i})) - \sum_{i=1}^{n}(1-y_i)\log(1+e^{\beta\cdot\mathbf{x}^i})$$
$$= \sum_{i=1}^{n} y_i \log(e^{\beta\cdot\mathbf{x}^i}) - \sum_{i=1}^{n}\log(1+e^{\beta\cdot\mathbf{x}^i})$$
$$= \sum_{i=1}^{n} y_i(\beta\cdot\mathbf{x}^i) - \sum_{i=1}^{n}\log(1+e^{\beta\cdot\mathbf{x}^i}).$$

ولتعظيم L نأخذ المشتقات المتعلقة بـ β_j. حيث أن:

$$\frac{d}{dz}\log(f(z)) = \frac{f'(z)}{f(z)} \text{ and } \frac{d}{dz}e^{za+b} = ae^{za+b},$$

ونحصل على

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^{n} y_i x_j^i - \sum_{i=1}^{n}\frac{x_j^i e^{\beta\cdot\mathbf{x}^i}}{1+e^{\beta\cdot\mathbf{x}^i}},\ j=1,2,\ldots,m,$$
$$\frac{\partial L}{\partial \beta_0} = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\frac{e^{\beta\cdot\mathbf{x}^i}}{1+e^{\beta\cdot\mathbf{x}^i}}.$$

وقمنا بوضع جميع المشتقات لتساوي صفر وحل المعادلات لإيجاد الحد الأقصى.

$$\sum_{i=1}^{n} x_j^i \left(y_i - \frac{e^{\beta \cdot \mathbf{x}^i}}{1 + e^{\beta \cdot \mathbf{x}^i}} \right) = 0,\ j = 1, 2, \ldots, m,$$

$$\sum_{i=1}^{n} \left(y_i - \frac{e^{\beta \cdot \mathbf{x}^i}}{1 + e^{\beta \cdot \mathbf{x}^i}} \right) = 0.$$

يتعين حل هذه المعادلات غير الخطية m+1 لإيجاد قيم m+1، $\beta_0,\beta_1,\ldots,\beta_m$. وليس من الصعب حل المعادلات باستخدام بعض الطرق التكرارية (مثل نيوتن–رافسون). وهذا هو الحل الذي يوفره برنامج تنفيذ الانحدار اللوجستي: فإن تقييم قيم β تزيد من فرصة مراقبة النمط 1_s و0_s الذين تمت ملاحظتهم.

10-4-2 إنشاء وثيقة التقييم

توفر وثيقة التقييم وسيلة لحساب درجة الفرد، ولكن الدرجات نفسها ليست مهمة جدا، شريطة أن نتمكن من النظر فيها مرة أخرى لمعرفة احتمال التخلف عن السداد. ولقد وجدنا بالفعل وثيقة تقييم تترجم مباشرة إلى (احتمالات لوغاريتمية) مضروبة في 100 في الجدول 10-4. وبالنسبة للانحدار اللوجستي، لدينا:

$$\log(\text{odds}) = \beta \cdot \mathbf{x} = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n.$$

وبما أن القيم x_i إما 0 أو 1، فإن قيم β هي بالضبط ما نحتاج إليه للحصول على وثيقة التقييم. لذلك، ببساطة نقوم بأخذ تقديرات الانحدار β، ونضربها في 100 للحصول على وثيقة التقييم.

لاحظ أن وثيقة التقييم سوف تشمل كل فئة حتى لو كان الانحدار اللوجستي ينتج تقديرات فقط للمتغيرات التي لا تنطوي عليها المتغيرات الأخرى. وهكذا، على سبيل المثال، في وثيقة التقييم بايز في الجدول 10-4 هناك درجة 23 لمن 'لديه بطاقة ائتمان' و61 لمن 'لا يحمل بطاقة ائتمان ". ومن شأن الانحدار اللوجستي أن ينتج درجة لواحد فقط من هذه المتغيرات، والآخر سيكون صفرا. ومن الواضح أن إضافة نفس العدد إلى كل مجموعة من المتغيرات المستثناة وغير المتكافئة هي نفس إضافة عدد إلى المدة الثابتة، ويحصل

كل فرد مهما كان تصنيفه على نفس النتيجة الإضافية. وهذا يعني أننا يمكن أن نتلاعب بحيث تكون جميع الدرجات متساوية. وعادة ما يتم ذلك من أجل تجنب أي أرقام سلبية في وثيقة التقييم.

لاحظ أنه بالنسبة لمعيار وثيقة التقييم (الاحتمالات اللوغاريتمية) المضروبة في 100، فإن زيادة قدرها 100 في الدرجات يعني زيادة قدرها 1 في الاحتمالات اللوغاريتمية، وهذا يتوافق مع ضرب الاحتمالات بالمبلغ الثابت e=2.7183. يمكننا أن ننظر إلى هذا بطريقة أخرى ونقول أن مضاعفة الاحتمالات بمقدار 2 سوف تزيد من النتيجة بمقدار 100 log(2)= 69. وفي بعض الأحيان نريد عمل وثيقة تقييم يوجد بها خاصية تطابق النتيجة الأساسية مع الاحتمالات المعطاة وتتطابق كل مضاعفة من الاحتمالات مع الزيادة في النتيجة بمقدار معين. للقيام بذلك نحن بحاجة إلى تعريف النتيجة كتحول خطي من الاحتمالات اللوغاريتمية. إذا أعطيت الدرجات لفئة i من خلال:

$$s_i = a + b\log(o_i)$$

بالنسبة للثوابت a وb، يمكن تحديد b من حقيقة أن مضاعفة الاحتمالات يتوافق مع زيادة الدرجات من خلال b log 2 ويمكن تحديد a من النتيجة الأساسية المطلوبة. وسنرى كيف يتم القيام بهذا من خلال مثال بنك سيدني أدناه.

مثال 10-1 (متابعة) بنك سيدني

استخدام نتائج الانحدار اللوجستي لعمل وثيقة تقييم لبيانات بنك سيدني حيث تكون جميع الأرقام إيجابية، وعلى أن تكون النتيجة مكونة من 200 درجة تقابل الاحتمالات من 10 إلى 1. وكل زيادة قدرها 50 في النتيجة تقابل ضرب الاحتمالات بعامل قدره 2.

الحل

نبدأ مع نتائج الانحدار اللوجستي للمعادلة (10-6). ويمكن مضاعفة القيم β بضربها في 100 وتقريبها لإعطاء وثيقة التقييم للاحتمالات اللوغاريتمية الأساسية مضروبة في 100

المسماه بوثيقة التقييم 1 في الجدول 10-8. بعد ذلك نقوم بإجراء تعديلات على وثيقة التقييم 1 من أجل الحصول على جميع الأرقام الإيجابية. ثم نأخذ مبلغ ثابت من المدة الثابتة وإضافته إلى كل فئة في مجموعة حصرية ومتبادلة.

لقد حققنا الحد الأدنى من التغيير لتحقيق ذلك: نقوم بطرح 135 من المدى الثابت وإضافة 98 إلى كل من المالك والمستأجر وغيرها، وبالإضافة إلى ذلك نضيف 37 لكل فئة عمرية. هذا من شأنه أن يعطينا وثيقة التقييم 2. لاحظ أن أي فرد سوف يحصل على نفس النتيجة بالضبط من وثيقة التقييم 2 كما تم الحصول عليها من وثيقة التقييم 1. الخطوة الأخيرة هي استخدام التحول الخطي من أجل تحقيق نتيجة مكونة من 200 درجة تقابل الاحتمالات من 10 إلى 1. وزيادة قدرها 50 عندما يتم مضاعفة الاحتمالات، سوف تحتوي وثيقة التقييم الجديدة على s=a+b log(odds) بدلاً من s=100 log(odds). (ترمز log(odds) إلى الاحتمالات اللوغاريتمية).

$$a + b\log(10) = 200$$
$$b\log(2) = 50.$$

وبالتالي $b = 50/0.693 = 72.135$. وبالتالي:

$$a = 200 - 72.135\log(10) = 33.903.$$

ولترجمة وثيقة التقييم للاحتمالات اللوغاريتمية المضروبة في 100 نحتاج أولا إلى مضاعفة جميع المتغيرات (بما في ذلك الثابت) بنسبة 72.135 / 100 ثم إضافة 33.9 إلى المتغير الثابت وتقريبه، ولكن لتطبيق وثيقة التقييم بشكل أسهل ، نقوم بخطوة إضافية لإضافة المتغير الثابت، الذي هو الآن 37×0.721+33.9=60.59، كل فئة من الفئات العمرية حتى نتمكن من إسقاط المتغير الثابت. وبعد التقريب، نحصل على وثيقة التقييم 3 كما هو مبين في الجدول 10-8.

الجدول 10-8: تعديل وثيقة التقييم من أجل الحصول على الخصائص المرغوبة.

السمات	وثيقة التقييم1	وثيقة التقييم 2	وثيقة التقييم 3
السن <30	0	37	87
السن 30–39	−37	0	61
السن 40–49	13	50	97
السن ≥50	20	57	102
يملك منزل	49	147	106
يستأجر منزل	−98	0	0
آخر	0	98	71
لديه بطاقة ائتمان	63	63	45
ليس لديه بطاقة ائتمان	0	0	0
ثابت	172	37	0

3-4-10 تطبيقات الدرجات الأخرى

بعد القيام بكل هذا العمل لتحديد إجراءات درجات الائتمان، فمن الجدير بالذكر أنه يمكن استخدام نفس التقنيات بالضبط لحل مشكلة إدارية أخرى شائعة. وهذا هو الهدف منها. فنحن لم نعد نريد أن نحدد الأشخاص الذين يرجح أن يكونوا سيئين من أجل تجنب منحهم المال، وبدلا من ذلك، نرغب في اختيار الأشخاص الذين يرجح أنهم لن يتخلفوا عن السداد وأيضاً استهدافهم. لا تزال فكرة نمذجة السلوك من حيث وثيقة التقييم قد تم خلقها من فئات مختلفة لا تزال قيمة. وتشترك كل من مشكلة تقييم درجات الائتمان ومشكلة استهداف الترويج في سمة أن نسبة صغيرة جدا من العينة هي في فئة الاستجابة (أو في فئة الديون المعدومة).

ومن وجهة نظر التقديرية، فإن النتيجة هي أن عددا صغيرا (مطلقا) من الأفراد في أي مجموعة معينة من الفئات سينتهي به المطاف إلى الاستجابة. هذا هو السبب في استخدام إجراء الانحدار اللوجستي غير المباشر بدلا من مجرد أخذ الاحتمالات التي نلاحظها في خانة واحدة في الجدول واستخدام هذا للتنبؤ بالاحتمالات اللوغاريتمية لهذه الخانة. وبسبب العدد الصغير من الأفراد المستجيبين في بعض الخانات، فإن التعامل مع كل خانة بمفردها هي طريقة سيئة للمضي قدما. ومن المرجح أن نحصل على نتيجة أفضل باستخدام نموذج الانحدار اللوجستي. ومن أجل تجنب النتيجة السلبية للاحتمالات

اللوغاريتمية ، فنحن بحاجة للتعامل مع احتمالات أكبر من 1. لذلك، في هذه المشكلة نقوم بتحديد الاحتمالات على النحو التالي:

$$\frac{\text{احتمالية عدم الاستجابة}}{\text{احتمالية الاستجابة}}.$$

عند تطبيق وثيقة التقييم نختار الأفراد الذين سنراسلهم عبر البريد بالقيم المنخفضة للنتيجة المقابلة لاحتمال عال نسبيا من الاستجابة. التمرين 10-5 هو مثال على هذا النوع من المشاكل.

ملاحظات

في مناقشة وكالات التصنيف الائتماني لقد استخدمت نطاق واسع من المعلومات التي تقدمها وكالة التصنيف الائتماني ستاندرد آند بورز. الجدول مأخوذ من الجدول 21. في دراسة ستاندرد أند بورز العالمية السنوية لتقييم وتقدير المراحل الانتقالية تاريخ النشر: 21 مارس 2012 (. ويوضح الشكل 10-2 التغيرات على مدار الوقت المبين في الجدول 3 من ذلك التقرير. وعلاوة على ذلك، يحتوي جدول البيانات BRMch10 Markov.xlsx على المواد المأخوذة من الجداول 21 و59 و61 و62 و64 من الدراسة الافتراضية السنوية للشركات العالمية في 2011 S & P. وتستخدم هذه المعلومات لمقارنة التنبؤات من نظرية ماركوف والسلوك الفعلي ينظر، الذي يشكل الأساس للشكلين 10-4 و10-5. الجدول الزمني لـ ليز كليبورن الوارد في الشكل 10-1 مأخوذ من المعلومات الواردة في الجدول 8 في ملخصات 2011 ستاندرد & بورز الافتراضية (تاريخ النشر: 21 مارس 2012).

هناك الكثير مما يمكن قوله عن الطريقة التي يمكن بها استخدام البيانات المحاسبية للاستدلال على مخاطر الائتمان. ويعود هذا العمل إلى نموذج معروف من قبل ألتمان في عام 1968، ويعطي مراجعة لهذه الأدبيات ألتمان وسوندرز (1998). استخدام تقنيات مثل الانحدار اللوجستي في هذا السياق يوفر صلة بين مخاطر الائتمان على مستوى الشركات والمستهلك. وهناك طرق أخرى في تقييم مخاطر الشركات وتعتمد على رؤية قيمة الشركة التي تتطور وفقا لعملية عشوائية. ثم، يمكن ترجمة المعلومات عن تقلب العملية والانحراف التصاعدي في القيمة إلى بيان حول احتمال التخلف عن السداد في

فترة T سنوات. وقد اقترحت مورتون في عام 2002 هذا النهج ومنذ ذلك الحين تم تبنيه من قبل شركة KMV، التي استحوذت عليها وكالة موديز في عام 2002. لمزيد من المعلومات حول هذا النهج، انظر الفصل 16 في Culp (2001) أو البحث الذي كتبه بهاراث وشومواي (2008).

تستند البيانات المستخدمة في مثال بنك سيدني، إلى مجموعة مماثلة من البيانات الافتراضية المبسطة التي قدمها توماس (2009) (وهو يسميها بيانات بنك ساوثامبتون). يعالج كتاب توماس بشكل أكثر تفصيلًا بكثير طريقة عمل نماذج الائتمان الاستهلاكية ويعد نقطة بداية جيدة إذا أردت التعمق في هذا الموضوع. في المناقشة نذكر مشاكل التربيع العادي اللوغريتمات إذا كانت الفئات تساوي صفر منفردة. وقد يكون من الناحية التفاؤلية جدًا أن نقول إن الانحدار اللوجستي يتجنب هذه المشكلة. هناك الكثير من الأعمال حول الطريقة التي يتم فيها التحيز للتقديرات القصوى لاحتمال الانحدار اللوجستي للعينات الصغيرة (انظر، على سبيل المثال، فيرث، 1993 والمراجع الواردة هناك). نهج واحد بسيط للحد من هذا التحيز هو إضافة 0.5 إلى كل من الخلايا "جيدة" و"سيئة". البرنامج غريتل المستخدم للانحدار اللوجستي متاح في http://gretl.sourceforge.net.

المراجع

Altman, E. and Saunders, A. (1998) Credit risk measurement: developments over the last 20 years. *Journal of Banking and Finance*, 21, 1721–1742.

Bharath, S. and Shumway, T. (2008) Forecasting default with the Merton distance to default model. *Review of Financial Studies*, 21, 1339–1369.

Culp, C. (2001) *The Risk Management Process: Business strategy and tactics*. John Wiley & Sons.

Firth, D. (1993) Bias reduction of maximum likelihood estimates. *Biometrika*, 80, 27–38.

Merton, R. (1974) On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29, 449–470.

Thomas, L. (2009) *Consumer Credit Models*. Oxford University Press, Oxford.

تمارين

1-10 إفلاس خلال عامين أو احتمالية ليست ذات صلة

وباستخدام المعلومات الواردة في الشكل 10-3 (بدلا من الأرقام المفصلة في الجدول 10-1)، احسب احتمال أن تكون الشركة التي صنفت على أنها "C /CCC" قد تم تصنيفها أنها متوقفة عن السداد أو مفلسة في خلال عامين.

2-10 مجموعات ماركوف

تحقق من الادعاء بأن مجموعة ولايات معا يمكن أن تلغي افتراض ماركوف. افترض أن هناك أربع ولايات A، B، C، D. وبمجرد الوصول إلى A أو D، لا يوجد أي تغيير ممكن. من B هناك فرصة 10٪ للانتقال إلى A و20٪ فرصة للانتقال إلى C، وإلا لا يوجد أي تغيير. وبالمثل، من C هناك فرصة 10٪ للانتقال إلى B و20٪ فرصة للانتقال إلى D، وإلا لا يوجد أي تغيير. الشركات الجديدة تصل إلى B في مثل هذه الطريقة التي نتوقع نفس العدد من الشركات في B كما أن هناك شركات في C. حساب احتمال ثلاث سنوات للوصول إلى D مع العلم أننا من المرجح أيضا أن تبدأ في إما B أو C، وقارن هذا مع التقدير الذي تم إجراؤه إذا جمعنا معا الولايتين B وC.

3-10 فئات ماركوف

التحقق من الادعاء بأن فئات شركات مختلفة ، كل منهم تتبع سلسلة ماركوف، يمكن أن تنتج سلوك غير ماركوف في المجموع. افترض أن هناك أربع ولايات A، B، C، D. وبمجرد الوصول إلى A أو D، لا يوجد أي تغيير ممكن. تتصرف الشركات من النوع X كما هو الحال في التمرين 10-2، أي بدءا من B، بعد سنة واحدة هناك فرصة 10٪ للانتقال إلى A و20٪ فرصة للانتقال إلى C. وبالمثل، من C هناك فرصة 10٪ للانتقال إلى B و20٪ من فرص الانتقال إلى D. الشركات من فئة Y هي نفسها إلا أنها تغير الولاية مرتين في كثير من الأحيان، أي من B هناك فرصة 20٪ للانتقال إلى A و40٪ فرصة

للانتقال إلى C. وبالمثل ، من C هناك فرصة 20٪ للانتقال إلى B و40٪ فرصة للانتقال إلى D. الشركات الجديدة تصل إلى B في مثل هذه الطريقة التي نتوقع أن يكون N الشركات من نوع X في B، N الشركات من نوع Y في شركات B وN من النوع X في C وN من النوع Y في C. احسب احتمال الانتقال إلى D في خطوتين من B لسلسلة ماركوف التي تتطابق مع التحولات السنوية الملحوظة ومقارنتها مع الاحتمال الحقيقي.

4-10 اكتوفون

اكتوفون هي شركة محمول تبقي بيانات عملائها وتقيمهم بناء على ما إذا فشلوا في سداد التزامات العقد أو لا في النصف الأول من العام المذكور في العقد. البيانات المتاحة على التطبيق ترتبط بالفئة العمرية، ما إذا كان لديهم بطاقة ائتمان وما إذا كان لديهم عقد للهاتف المحمول (إما مع أوكتوفون أو شركة أخرى). وترد النتائج من 2000 من المتقدمين في العام الماضي الذين يعيشون في مدينة واحدة في الجدول 10-9. في هذا الجدول هناك 1850 نتيجة جيدة و150 سيئة. حساب الأوزان من الأدلة لمختلف السمات المعنية واستخدامها لإنشاء (naive Bayes ') بطاقة الأداء.

الجدول 10-3: بيانات اكتوفون: النتائج الجيدة والسيئة

	العمر 18–21	العمر 22–29	العمر 30–45	46 فيما فوق
المحمول السابق، بطاقة الائتمان	$G = 150$ $B = 6$	$G = 256$ $B = 8$	$G = 312$ $B = 9$	$G = 250$ $B = 7$
المحمول السابق، بدون بطاقة ائتمان	$G = 114$ $B = 10$	$G = 123$ $B = 13$	$G = 92$ $B = 9$	$G = 91$ $B = 12$
لا يوجد محمول سابق، يوجد بطاقة ائتمان	$G = 99$ $B = 12$	$G = 182$ $B = 6$	$G = 59$ $B = 11$	$G = 45$ $B = 8$
لا يوجد محمول سابق ولا يوجد بطاقة ائتمان	$G = 22$ $B = 9$	$G = 26$ $B = 14$	$G = 13$ $B = 7$	$G = 16$ $B = 9$

5-10 اكتوفون مع تكاليف العقود

تبيع أوكتوفون مجموعة متنوعة من العقود ولكن يمكن تصنيفها على أساس دولار

في الشهر إلى ثلاث فئات: منخفضة التكلفة – أقل من 30 دولار، تكلفة المتوسطة – ما بين 30 و40 دولار؛ وتكلفة مرتفعة – أكثر من 40 دولار.

(أ) من 2000 في عينة تمرين 10-4 هناك 800 مجموعه من العقود منخفضة التكلفة، و600 عقد متوسط التكلفة و600 عقد مرتفع التكلفة. ويتم توزيع 150 من العقود السيئة مع 40 على عقود منخفضة التكلفة، و40 على العقود متوسطة التكلفة و70 على العقود عالية التكلفة. احسب الأوزان الجديدة للأدلة بها في ذلك هذه المعلومات الإضافية.

(ب) ويتم إجراء انحدار لوجستي على هذه البيانات وينتج المعاملات المبينة في الجدول 10-10. هنا، سمات عمر 45+؛ بدون بطاقة ائتمان؛ ولا محمول سابق؛ والعقود عالية التكلفة كلها تم حذفها بسبب الخطية المتداخلة. احسب بطاقة الأداء باستخدام هذه البيانات.

الجدول 10-10: انحدار لوجيستي لاكتوفون

	المعامل
const	0.194171
age18_21	0.135372
age22_29	0.401779
age30_45	0.00415712
credit_card	1.38715
prev_phone	1.45171
low_cost	0.913791
med_cost	0.427606

(ج) استخدام التحجيم لضبط بطاقة الأداء من (ب) بحيث يكون لها الخصائص التالية: (أ) ليس هناك مصطلح ثابت. (ب) الأداء بدرجة 500 تمثل احتمالات تتراوح بين 100 و1؛ و(ج) زيادة في درجة تحجيم 100 نقطة تمثل احتمالات يتم ضربها في 10.

(حتى الأداء بدرجة 600 يمثل خلاف 1000 إلى 1.)

(د) تم احتساب عقد الهاتف المحمول بتحقيقه ربحًا متوسطا قدره 15 دولارا في السنة، في حين أن كل عميل يفشل في سداد مدفوعات مجدولة سيكلف في المتوسط 100 دولار (بما في ذلك تكلفة المتابعة ومتوسط الديون غير المحصلة وقيمة الاستبدال تلك الهواتف التي لم يتم استردادها، بعد السماح لمتوسط المدفوعات التي تم إجراؤها بالفعل). لبطاقة النتائج في (ج)، ما هي النتيجة المستقطعة التي تجعل اوكتوفون تفسخ العقد؟

(هـ) لنفترض بأن النظر في الأرقام بشكل أكثر تفصيلًا يشير إلى أن الأرباح والخسائر تعتمد على حجم العقد على النحو التالي:

	تكلفة منخفضة	تكلفة متوسطة	تكلفة مرتفعة
الربح السنوي	$10	$15	$20
متوسط التكلفة إذا كان سيئًا	$80	$100	$160

بماذا توصي اكتوفون؟

6-10 Cosal Skincare لمنتجات الرعاية بالبشرة

تبيع شركةCosal للعناية بالبشرة منتجات التجميل عبر الإنترنت. كل ستة أشهر تبدأ Cosal حملة للإعلان عن أحد منتجاتها الجديدة في محاولة لزيادة قاعدة عملائها، من خلال إرسال عينة مجانية عبر البريد للعملاء المحتملين على قائمة بريدية. تتضمن القائمة البريدية، التي تشتريها Cosal ، سمات معينة للمستلمين. على سبيل المثال، يمكن لـ Cosal اختيار الإناث تحت سن 30 فقط الذين لديهم رخصة قيادة ويعيشون وحدهم. يستخدم كوزال البيانات من الماضي تعزيز لتوجيه ذلك لنوع من العملاء المحتملين لتشمل في البريد بها. كما هو شائع مع هذا النوع التسويق، هناك نسبة صغيرة من المتلقين للعينة المجانية سيشترون المنتج عبر الإنترنت. ومع ذلك، ترى Cosal عملية شراء واحدة على الإنترنت يتبعها عمليات شراء أخرى. وتشير التقديرات إلى أن الأرباح المكتسبة من عميل واحد جديد عبر الإنترنت تبلغ حوالي 150 دولارا في المتوسط. وتكلف العينة المجانية 2.50 دولار

لكل متلقي بما في ذلك تكاليف البريد.

الجدول 10-11 يوضح النتائج من الرسائل البريدية المرسلة لـ 10000 عميل محتمل. من خلال هذه الطريقة، تم تحقيق 310 عميل جديد على الإنترنت، وهي نتيجة جيدة للشركة. في هذا الجدول، نتائج P هي أولئك الذين اشتروا على الإنترنت، ونتائج N أولئك الذين لم يجرون عملية شراء عبر الإنترنت.

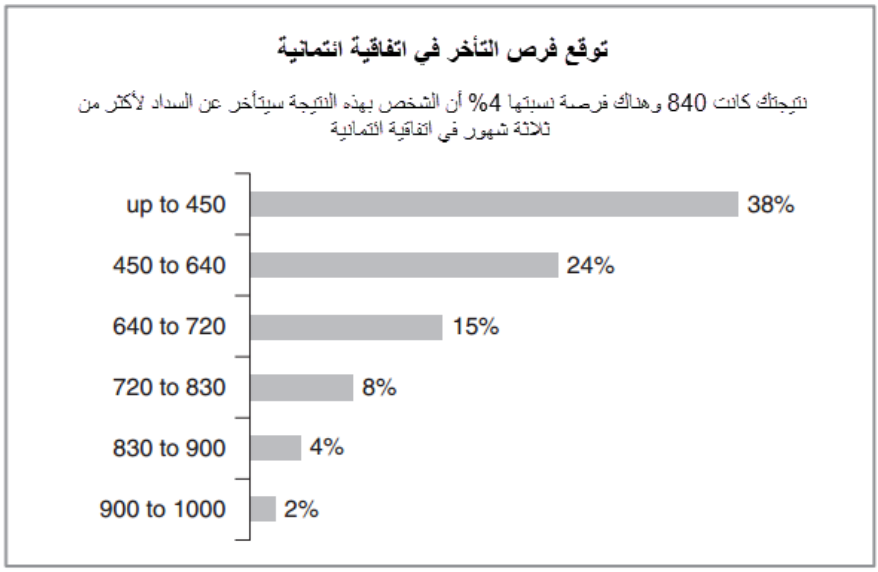
(أ) استخدم هذه البيانات (متوفرة أيضا في جدول بيانات على ملف إكسل BRMch10 CosalSkincare.xls) للعمل على فرص اللوغاريتمات لكل فئة ومن طبق مربعات صغرى عادية أقل انحدار لتحديد مجموعة من القيم β للاستخدام.

	تحت 30	30–39	40–49	فوق 50
أسرة 1، بوظيفة بداو كامل	$N = 456$	$N = 412$	$N = 386$	$N = 386$
	$P = 34$	$P = 16$	$P = 8$	$P = 15$
أسرة 1، بوظيفة ليست بدوام كامل	$N = 453$	$N = 389$	$N = 375$	$N = 395$
	$P = 24$	$P = 12$	$P = 6$	$P = 12$
أسرة 2، بوظيفة بداو كامل	$N = 463$	$N = 409$	$N = 387$	$N = 363$
	$P = 24$	$P = 15$	$P = 10$	$P = 15$
أسرة 2، بوظيفة ليست بدوام كامل	$N = 460$	$N = 401$	$N = 373$	$N = 371$
	$P = 18$	$P = 13$	$P = 5$	$P = 18$
أسرة > 3، بوظيفة بدوام كامل	$N = 448$	$N = 407$	$N = 405$	$N = 375$
	$P = 12$	$P = 6$	$P = 6$	$P = 9$
أسرة > 3، بوظيفة ليست بدوام كامل	$N = 412$	$N = 404$	$N = 377$	$N = 383$
	$P = 12$	$P = 8$	$P = 4$	$P = 8$

(ب) ضع أهداف لتمكين Cosal Skincare من اتخاذ قرار بشأن الأفراد الذي يجب أن يتم إدراجهم في القائمة البريدية.

7-10 الاتساق مع المربعات الصغرى العادية

لنفترض أن الشكل 10-7 يتم إنتاجه من قبل موقع على شبكة الإنترنت يقدم تقييم درجة الائتمان مجانا. هل الأرقام تظهر اتساقًا مع المربعات الصغرى العادية لمخطط الأداء؟



الشكل 10-7: رسم من موقع يقدم تصنيفات ائتمانية مجانا.