

Ασκήσεις Πιστωτικού Κινδύνου

Άσκηση 1: Υπόδειγμα Logit για PD, Απόφαση Πίστωσης και Αναμενόμενη Ζημία

Ερώτημα

Μια τράπεζα εκτιμά το ακόλουθο υπόδειγμα Logit για την πιθανότητα αθέτησης εταιρικών δανειοληπτών:

$$Z_i = -3.20 - 0.85ICR_i + 1.10DE_i - 0.60PM_i + 0.75Late_i - 0.40CQ_i$$

$$PD_i = \frac{1}{1 + e^{-Z_i}}$$

όπου:

Μεταβλητή	Ερμηνεία
<i>ICR</i>	Δείκτης κάλυψης τόκων (Interest Coverage Ratio)
<i>DE</i>	Δείκτης ξένων προς ίδια κεφάλαια (Debt-to-Equity Ratio)
<i>PM</i>	Περιθώριο κέρδους (Profit Margin)
<i>Late</i>	Ψευδομεταβλητή ίση με 1 εάν ο δανειολήπτης είχε καθυστερήσεις πληρωμών
<i>CQ</i>	Ψευδομεταβλητή ίση με 1 εάν η ποιότητα των εξασφαλίσεων είναι ισχυρή

Η τράπεζα εξετάζει τις ακόλουθες αιτήσεις δανείων:

Δανειολήπτης	<i>ICR</i>	<i>DE</i>	<i>PM</i>	<i>Late</i>	<i>CQ</i>	<i>EAD</i> (€)	<i>LGD</i>
A	3.0	0.8	0.12	0	1	1.000.000	35%
B	1.2	2.0	0.03	1	0	700.000	55%
C	2.0	1.4	0.06	0	0	500.000	45%

Η τράπεζα εγκρίνει ένα δάνειο μόνο εάν:

$$PD_i < 5\%$$

Να απαντήσετε στα ακόλουθα:

- Να υπολογιστούν τα Z_i και PD_i για κάθε δανειολήπτη.
- Να αποφασιστεί αν κάθε δάνειο εγκρίνεται ή απορρίπτεται.
- Να υπολογιστεί η Αναμενόμενη Ζημία (Expected Loss):

$$EL_i = PD_i \times LGD_i \times EAD_i$$

- Να δοθεί σύντομη ερμηνεία των αποτελεσμάτων.

Ενδεικτική απάντηση

Δανειολήπτης A

$$Z_A = -3.20 - 0.85(3.0) + 1.10(0.8) - 0.60(0.12) + 0.75(0) - 0.40(1)$$

$$Z_A = -5.342$$

$$PD_A = \frac{1}{1 + e^{5.342}} = 0.00476 = 0.48\%$$

$$EL_A = 0.00476 \times 0.35 \times 1,000,000 = 1,666$$

Δανειολήπτης B

$$Z_B = -3.20 - 0.85(1.2) + 1.10(2.0) - 0.60(0.03) + 0.75(1) - 0.40(0)$$

$$Z_B = -1.288$$

$$PD_B = \frac{1}{1 + e^{1.288}} = 0.2162 = 21.62\%$$

$$EL_B = 0.2162 \times 0.55 \times 700,000 = 83,237$$

Δανειολήπτης C

$$Z_C = -3.20 - 0.85(2.0) + 1.10(1.4) - 0.60(0.06) + 0.75(0) - 0.40(0)$$

$$Z_C = -3.396$$

$$PD_C = \frac{1}{1 + e^{3.396}} = 0.0324 = 3.24\%$$

$$EL_C = 0.0324 \times 0.45 \times 500,000 = 7,290$$

Συνοπτικά αποτελέσματα

Δανειολήπτης	Z_i	PD	Απόφαση	EL (€)
A	-5.342	0.48%	Έγκριση	1.666
B	-1.288	21.62%	Απόρριψη	83.237
C	-3.396	3.24%	Έγκριση	7.290

Ερμηνεία

Ο δανειολήπτης B πρέπει να απορριφθεί, καθώς η εκτιμώμενη πιθανότητα αθέτησης υπερβαίνει σημαντικά το όριο έγκρισης του 5%. Η υψηλή PD συνδέεται με χαμηλή κάλυψη τόκων, υψηλή μόχλευση, χαμηλή κερδοφορία, προηγούμενες καθυστερήσεις πληρωμών και ασθενέστερη ποιότητα εξασφαλίσεων. Επιπλέον, ο δανειολήπτης B εμφανίζει τη μεγαλύτερη αναμενόμενη ζημία. Το παράδειγμα δείχνει ότι ο πιστωτικός κίνδυνος δεν εξαρτάται μόνο από την πιθανότητα αθέτησης, αλλά από τον συνδυασμό PD, LGD και EAD.

Άσκηση 2: Μονομεταβλητό Scorecard, Επιλογή Διαστημάτων βάσει IV και Απόφαση Έγκρισης/Απόρριψης

Ερώτημα

Μια τράπεζα θέλει να κατασκευάσει ένα μονομεταβλητό credit scorecard χρησιμοποιώντας μόνο τη μεταβλητή **Ηλικία**.

Ο αναλυτής εξετάζει δύο εναλλακτικές ταξινομήσεις της ηλικίας σε διαστήματα. Στόχος είναι να επιλεγεί η ταξινόμηση με το υψηλότερο Information Value (IV) και στη συνέχεια να χρησιμοποιηθούν τα επιλεγμένα διαστήματα σε έναν κανόνα πιστωτικής απόφασης βασισμένο σε υπόδειγμα Logit.

Οι τύποι είναι:

$$WoE_j = \ln \left(\frac{\%NonEvents_j}{\%Events_j} \right)$$

$$IV = \sum_j (\%NonEvents_j - \%Events_j) \times WoE_j$$

όπου:

- **Non-event:** ο δανειολήπτης δεν αθέτησε τις υποχρεώσεις του.
- **Event:** ο δανειολήπτης αθέτησε τις υποχρεώσεις του.

Ταξινόμηση A: 2 διαστήματα

Διάστημα ηλικίας	Non-events	Events
Έως 40	380	50
41 και άνω	520	50
Σύνολο	900	100

Ταξινόμηση B: 4 διαστήματα

Διάστημα ηλικίας	Non-events	Events
Έως 25	80	20
26-40	300	30
41-65	360	35
66 και άνω	160	15
Σύνολο	900	100

Μετά την επιλογή της καλύτερης ταξινόμησης, η τράπεζα εκτιμά το ακόλουθο μονομεταβλητό υπόδειγμα Logit:

$$Z_i = -2.20 - 1.20WoE(Age_i)$$

$$PD_i = \frac{1}{1 + e^{-Z_i}}$$

Η τράπεζα εφαρμόζει τον ακόλουθο κανόνα πιστωτικής απόφασης:

Έγκριση εάν $PD_i < 10\%$

Απόρριψη εάν $PD_i \geq 10\%$

Να αξιολογηθούν οι ακόλουθοι πελάτες:

Πελάτης	Ηλικία
A	22
B	35
C	50
D	70

Να απαντήσετε στα ακόλουθα:

1. Να υπολογιστούν το WoE και το IV για την Ταξινόμηση A.
2. Να υπολογιστούν το WoE και το IV για την Ταξινόμηση B.
3. Να επιλεγεί η καλύτερη ταξινόμηση των διαστημάτων.
4. Να εκτιμηθεί η PD για κάθε επιλεγμένο διάστημα ηλικίας.
5. Να αποφασιστεί αν κάθε πελάτης εγκρίνεται ή απορρίπτεται.

Ενδεικτική απάντηση

Βήμα 1: IV για την Ταξινόμηση A — 2 διαστήματα

Σύνολο Non-events = 900

Σύνολο Events = 100

Διάστημα ηλικίας	% Non-events	% Events	WoE	Συνιστώσα IV
Έως 40	$380/900 = 42.22\%$	$50/100 = 50.00\%$	$\ln(0.4222/0.5000) = -0.169$	0.013
41 και άνω	$520/900 = 57.78\%$	$50/100 = 50.00\%$	$\ln(0.5778/0.5000) = 0.145$	0.011

$$IV_A = 0.013 + 0.011 = 0.024$$

Η ταξινόμηση Α έχει ασθενή προβλεπτική ικανότητα.

Βήμα 2: IV για την Ταξινόμηση Β — 4 διαστήματα

Διάστημα ηλικίας	% Non-events	% Events	WoE	Συνιστώσα IV
Έως 25	80/900 = 8.89%	20/100 = 20.00%	$\ln(0.0889/0.2000) = -0.811$	0.090
26-40	300/900 = 33.33%	30/100 = 30.00%	$\ln(0.3333/0.3000) = 0.105$	0.004
41-65	360/900 = 40.00%	35/100 = 35.00%	$\ln(0.4000/0.3500) = 0.134$	0.007
66 και άνω	160/900 = 17.78%	15/100 = 15.00%	$\ln(0.1778/0.1500) = 0.170$	0.005

$$IV_B = 0.090 + 0.004 + 0.007 + 0.005 = 0.105$$

Βήμα 3: Επιλογή ταξινόμησης

Ταξινόμηση	IV	Ερμηνεία
A — 2 διαστήματα	0.024	Ασθενής προβλεπτική ικανότητα
B — 4 διαστήματα	0.105	Υψηλότερη/καλύτερη προβλεπτική ικανότητα

Επιλεγμένη ταξινόμηση = Ταξινόμηση Β

Άρα το scorecard χρησιμοποιεί τα ακόλουθα διαστήματα ηλικίας: Έως 25, 26-40, 41-65 και 66 και άνω.

Βήμα 4: Υπολογισμός PD ανά επιλεγμένο διάστημα ηλικίας

$$Z_i = -2.20 - 1.20WoE(Age_i)$$

$$PD_i = \frac{1}{1 + e^{-Z_i}}$$

Διάστημα ηλικίας: Έως 25

$$Z = -2.20 - 1.20(-0.811) = -1.227$$

$$PD = \frac{1}{1 + e^{1.227}} = 22.67\%$$

Διάστημα ηλικίας: 26-40

$$Z = -2.20 - 1.20(0.105) = -2.326$$

$$PD = \frac{1}{1 + e^{2.326}} = 8.90\%$$

Διάστημα ηλικίας: 41-65

$$Z = -2.20 - 1.20(0.134) = -2.361$$

$$PD = \frac{1}{1 + e^{2.361}} = 8.63\%$$

Διάστημα ηλικίας: 66 και άνω

$$Z = -2.20 - 1.20(0.170) = -2.404$$

$$PD = \frac{1}{1 + e^{2.404}} = 8.29\%$$

Βήμα 5: Πιστωτική απόφαση ανά διάστημα ηλικίας

Διάστημα ηλικίας	WoE	Z_i	PD	Απόφαση
Έως 25	-0.811	-1.227	22.67%	Απόρριψη
26-40	0.105	-2.326	8.90%	Έγκριση
41-65	0.134	-2.361	8.63%	Έγκριση
66 και άνω	0.170	-2.404	8.29%	Έγκριση

Βήμα 6: Αποφάσεις σε επίπεδο πελάτη

Πελάτης	Ηλικία	Διάστημα ηλικίας	WoE	PD	Απόφαση
A	22	Έως 25	-0.811	22.67%	Απόρριψη
B	35	26-40	0.105	8.90%	Έγκριση
C	50	41-65	0.134	8.63%	Έγκριση
D	70	66 και άνω	0.170	8.29%	Έγκριση

Ερμηνεία

Η ταξινόμηση τεσσάρων διαστημάτων επιλέγεται επειδή παράγει υψηλότερο IV από την ταξινόμηση δύο διαστημάτων. Η νεότερη ηλικιακή ομάδα έχει αρνητικό WoE, καθώς το ποσοστό των αθετήσεων είναι υψηλό σε σχέση με το ποσοστό των μη αθετήσεων. Συνεπώς, λαμβάνει υψηλότερη εκτιμώμενη PD και απορρίπτεται βάσει του ορίου του 10%. Τα υπόλοιπα διαστήματα έχουν θετικό WoE, χαμηλότερη εκτιμώμενη PD και εγκρίνονται.

Άσκηση 3: Αξιολόγηση Υποδείγματος, Πίνακες Σύγχυσης και Σημεία ROC

Ερώτημα

Μια τράπεζα αξιολογεί ένα υπόδειγμα PD σε δείγμα 2.000 δανειοληπτών. Οι δανειολήπτες ομαδοποιούνται με βάση την εκτιμώμενη πιθανότητα αθέτησης.

Ζώνη PD	Αριθμός δανειοληπτών	Πραγματικές αθετήσεις	Πραγματικές μη αθετήσεις
$PD \geq 10\%$	120	42	78
$8\% \leq PD < 10\%$	160	30	130
$5\% \leq PD < 8\%$	260	28	232
$3\% \leq PD < 5\%$	360	20	340
$1\% \leq PD < 3\%$	600	15	585
$PD < 1\%$	500	5	495
Σύνολο	2.000	140	1.860

Η τράπεζα αξιολογεί το υπόδειγμα χρησιμοποιώντας τα ακόλουθα σημεία cut-off:

10%, 8%, 5%, 3%, 1%

Για κάθε σημείο cut-off, ένας δανειολήπτης ταξινομείται ως **Predicted Bad** εάν:

$$PD \geq \text{cut-off}$$

και ως **Predicted Good** εάν:

$$PD < \text{cut-off}$$

Να απαντήσετε στα ακόλουθα:

- Να κατασκευαστεί ο πίνακας σύγχυσης (confusion matrix) για κάθε σημείο cut-off: True Positive (TP), False Positive (FP), False Negative (FN) και True Negative (TN).
- Να υπολογιστούν τα TPR και FPR :

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

3. Να κατασκευαστούν τα σημεία της καμπύλης ROC.
4. Να παρουσιαστεί ένα απλό γράφημα ROC.

Ενδεικτική απάντηση

Βήμα 1: Σύνολο αθετήσεων και μη αθετήσεων

Σύνολο πραγματικών αθετήσεων = 140

Σύνολο πραγματικών μη αθετήσεων = 1,860

Επομένως:

$$TP + FN = 140$$

$$FP + TN = 1,860$$

Βήμα 2: Πίνακας σύγκρισης για κάθε σημείο cut-off

Σημείο cut-off = 10%

Predicted Bad:

$$PD \geq 10\%$$

	Actual Bad	Actual Good
Predicted Bad	$TP = 42$	$FP = 78$
Predicted Good	$FN = 98$	$TN = 1,782$

Σημείο cut-off = 8%

Predicted Bad:

$$PD \geq 8\%$$

	Actual Bad	Actual Good
Predicted Bad	$TP = 72$	$FP = 208$
Predicted Good	$FN = 68$	$TN = 1,652$

Σημείο cut-off = 5%

Predicted Bad:

$$PD \geq 5\%$$

	Actual Bad	Actual Good
Predicted Bad	$TP = 100$	$FP = 440$
Predicted Good	$FN = 40$	$TN = 1,420$

Σημείο cut-off = 3%

Predicted Bad:

$$PD \geq 3\%$$

	Actual Bad	Actual Good
Predicted Bad	$TP = 120$	$FP = 780$
Predicted Good	$FN = 20$	$TN = 1,080$

Σημείο cut-off = 1%

Predicted Bad:

$$PD \geq 1\%$$

	Actual Bad	Actual Good
Predicted Bad	$TP = 135$	$FP = 1,365$
Predicted Good	$FN = 5$	$TN = 495$

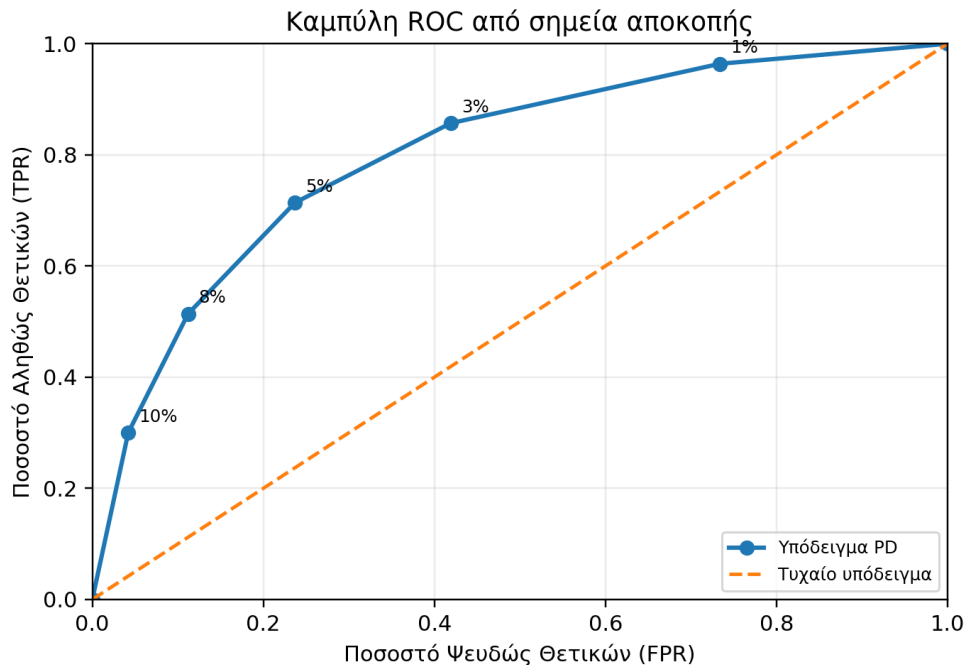
Βήμα 3: Σύνοψη σημείων ROC

Κάθε σημείο cut-off δίνει ένα σημείο στην καμπύλη ROC:

(FPR, TPR)

Cut-off	TP	FP	FN	TN	TPR	FPR
Start	0	0	140	1.860	0.000	0.000
10%	42	78	98	1.782	0.300	0.042
8%	72	208	68	1.652	0.514	0.112
5%	100	440	40	1.420	0.714	0.237
3%	120	780	20	1.080	0.857	0.419
1%	135	1.365	5	495	0.964	0.734
All borrowers	140	1.860	0	0	1.000	1.000

Βήμα 4: Γράφημα ROC



Καμπύλη ROC από σημεία αποκοπής (cut-off)

Η καμπύλη ROC απεικονίζει το ποσοστό αληθώς θετικών (True Positive Rate - TPR) έναντι του ποσοστού ψευδώς θετικών (False Positive Rate - FPR) για διαφορετικά σημεία cut-off. Ένα υπόδειγμα χωρίς διακριτική ικανότητα θα βρισκόταν κοντά στη διαγώνια γραμμή, ενώ ένα καλύτερο υπόδειγμα βρίσκεται πάνω από αυτήν.