

DÉPARTEMENT MATHÉMATIQUES

EXAMEN DE STATISTIQUE II

06 février 2008 Durée 2h00 heures

Consignes

- Toutes vos réponses doivent être justifiées
- *La consultation des documents et calculatrices est interdite.*
Seul document autorisé : Deux feuilles format A4, manuscrites recto-verso.
- L'utilisation de la grille fournie n'est pas obligatoire mais elle peut faciliter votre travail.

Exercice 1 .- Dans un atelier une machine fabrique des pièces en grande série; on s'intéresse à leur longueur mesurée en cm. On admet que la v.a. qui, à chaque pièce tirée au hasard associe sa longueur, suit une loi normale de moyenne $m = 150$. Afin de contrôler que la machine est toujours bien réglée, on se propose d'effectuer un test d'hypothèse, avec $\alpha = 10\%$. On prélève des échantillons de 49 pièces (tirage aléatoire avec remise). À chaque échantillon, on associe la moyenne des longueurs des 49 pièces.

1. Construire les hypothèses de ce test.
2. Représenter graphiquement ce test.
3. Quelle est la loi de la v.a. \bar{X} sous l'hypothèse H_0 ?
4. Calculer la région critique
5. Peut-on calculer le risque β ?
6. Énoncer la règle de décision permettant d'accepter si la machine est toujours bien réglée.
7. La moyenne observée sur un échantillon de 49 pièces est $\bar{x} = 149,5$ et un écart-type 10. Que peut-on conclure quant à la qualité des pièces produites avec $\alpha = 10\%$?
8. Peut-on améliorer ce test et comment?
9. Indiquer comment varie la région critique en fonction de la taille n de l'échantillon.

Exercice 2 .- Chez Aluminax, on désire expliquer la résistance d'une fibre, appelée LIANTE, fabriquée à partir de quatre produits X_1, X_2, X_3, X_4 qui interviennent à sa fabrication.

- X_1 : mesuré en %
- X_2 : mesuré en grammes
- X_3 : mesuré en cm^3
- X_4 : mesuré en %

Pour des raisons aussi bien d'efficacité industrielle et de sécurité, on cherche à comprendre et à maîtriser la résistance de la fibre. On prend le risque $\alpha = 10\%$

1. Choix d'un modèle

- (a) Commenter le modèle $Y/X_1X_2X_3X_4$
- (b) Retenir le meilleur modèle, en utilisant la méthode pas à pas.
- (c) Commenter et expliquer votre choix.

2. On considère maintenant le modèle ci-dessous choisi au hasard:

$$Y = a_1X_1 + a_3X_3 + a_0 + \epsilon$$

- (a) Le responsable de la production prétend que s'il augmente de cinq volumes le produit X_3 , il augmentera de cinq unités la résistance de la fibre. Pouvez-vous confirmer ses dires?
- (b) Construire et répondre au test: " la variable X_1 joue un rôle significatif dans le modèle considéré, avec $\alpha = 10\%$ ".
- (c) Doit-on modifier ce modèle, si on mesure la variable $X_3 m^3$? Commenter
- (d) Quelle est la signification du coefficient de corrélation multiple?
- (e) Peut-il exister plusieurs droites de régression pour un échantillon donné?

Chimie : Analyse n° 1

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées 44
Nombre de variables utilisées 2
Variable à expliquer LIAN
RESULTATS
Coefficient de corrélation multiple R5502
Carré du coef de corrélation multiple R² .. .3027
Valeur du test F de signification de R² ... 18.2295 avec 1 et 43 ddl
Somme des carrés des écarts SCE 13.3092
Variance résiduelle3169

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 42 ddl
X1	-15.287333	3.580508	4.2695996
Terme constant	42.132394		

Chimie : Analyse n° 2

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées 44
Nombre de variables utilisées 2
Variable à expliquer LIAN
RESULTATS
Coefficient de corrélation multiple R4064
Carré du coef de corrélation multiple R² .. .1652
Valeur du test F de signification de R² ... 8.3096 avec 1 et 43 ddl
Somme des carrés des écarts SCE 15.9334
Variance résiduelle3794

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 42 ddl
X2	.334990	.116209	2.8826436
Terme constant	22.387750		

Chimie : Analyse n° 3

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées 44
Nombre de variables utilisées 2
Variable à expliquer LIAN

RESULTATS
Coefficient de corrélation multiple R3990
Carré du coef de corrélation multiple R² .. .1592
Valeur du test F de signification de R² ... 7.9513 avec 1 et 43 ddl
Somme des carrés des écarts SCE 16.0477
Variance résiduelle3821

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 42 ddl
X3	- .327160	.116023	2.8198006
Terme constant	63.778616		

Chimie : Analyse n° 4

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées 44
Nombre de variables utilisées 2
Variable à expliquer LIAN

RESULTATS
Coefficient de corrélation multiple R5033
Carré du coef de corrélation multiple R² .. .2533
Valeur du test F de signification de R² ... 14.2448 avec 1 et 43 ddl
Somme des carrés des écarts SCE 14.2520
Variance résiduelle3393

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif.	à	42 ddl
X4	.051736	.013708	3.7742356		
Terme constant	25.928633				

Chimie : Analyse n° 5 |

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées	44
Nombre de variables utilisées	3
Variable à expliquer	LIAN
2e variable explicative	X2

RESULTATS

Coefficient de corrélation multiple R7101
Carré du coef de corrélation multiple R^2 ..	.5042
Valeur du test F de signification de R^2 ...	20.8506 avec 2 et 42 ddl
Somme des carrés des écarts SCE	9.4620
Varianc e résiduelle2308

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif.	à	41 ddl
X1	-16.226324	3.064221	5.2954166		
X2	.371115	.090895	4.0829226		
Terme constant	35.578105				

Chimie : Analyse n° 6 |

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées	44
Nombre de variables utilisées	3
Variable à expliquer	LIAN
2e variable explicative	X3

RESULTATS

Coefficient de corrélation multiple R6340
Carré du coef de corrélation multiple R^2 ..	.4020
Valeur du test F de signification de R^2 ...	13.7781 avec 2 et 42 ddl
Somme des carrés des écarts SCE	11.4142
Varianc e résiduelle2784

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif.	à	41 ddl
X1	-13.869680	3.399735	4.0796366		
X3	- .261743	.100326	2.6089356		
Terme constant	68.717466				

Chimie : Analyse n° 7 |

===== REGRESSION MULTIPLE =====

Nombre d'observations utilisées	44
Nombre de variables utilisées	3
Variable à expliquer	LIAN
2e variable explicative	X4

RESULTATS

Coefficient de corrélation multiple R5966
Carré du coef de corrélation multiple R^2 ..	.3559
Valeur du test F de signification de R^2 ...	11.3290 avec 2 et 42 ddl
Somme des carrés des écarts SCE	12.2925
Varianc e résiduelle2998

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif.	à	41 ddl
X1	-10.837211	4.239097	2.5564906		
X4	.028879	.015683	1.8414156		
Terme constant	36.587006				

Chimie : Analyse n° 8 |

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées 44
Nombre de variables utilisées 3
Variable à expliquer LIAN
2e variable explicative X3

RESULTATS

Coefficient de corrélation multiple R4966
Carré du coef de corrélation multiple R^22466
Valeur du test F de signification de R^2 ... 6.7103 avec 2 et 42 ddl
Somme des carrés des écarts SCE 14.3791
Varianc e résiduelle3507

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 41 ddl
X2	.256824	.117742	2.1812446
X3	-.246589	.117133	2.1052056
Terme constant	50.128347		

Chimie : Analyse n° 9 |

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées 44
Nombre de variables utilisées 3
Variable à expliquer LIAN
2e variable explicative X4

RESULTATS

Coefficient de corrélation multiple R7168
Carré du coef de corrélation multiple R^25138
Valeur du test F de signification de R^2 ... 21.6597 avec 2 et 42 ddl
Somme des carrés des écarts SCE 9.2804
Varianc e résiduelle2264

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 41 ddl
X2	.428363	.091401	4.6866086
X4	.061803	.011400	5.4214796
Terme constant	16.830829		

Chimie : Analyse n° 10 |

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées 44
Nombre de variables utilisées 3
Variable à expliquer LIAN
2e variable explicative X4

RESULTATS

Coefficient de corrélation multiple R7401
Carré du coef de corrélation multiple R^25477
Valeur du test F de signification de R^2 ... 24.8271 avec 2 et 42 ddl
Somme des carrés des écarts SCE 8.6319
Varianc e résiduelle2105

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 41 ddl
X3	-.459687	.088971	5.1666956
X4	.066200	.011154	5.9349726
Terme constant	73.885442		

Chimie : Analyse n° 11 |

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées 44
Nombre de variables utilisées 4
Variable à expliquer LIAN
2e variable explicative X2

3e variable explicative X3

RESULTATS

Coefficient de corrélation multiple R7315
Carré du coef de corrélation multiple R ² ..	.5351
Valeur du test F de signification de R ² ...	15.3493 avec 3 et 41 ddl
Somme des carrés des écarts SCE	8.8722
Variance résiduelle2218

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 40 ddl
X1	-15.258019	3.062163	4.9827596
X2	.319876	.094487	3.3853966
X3	-.154843	.094954	1.6307126
Terme constant	52.210330		

Chimie : Analyse n° 12

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées	44
Nombre de variables utilisées	4
Variable à expliquer	LIAN
2e variable explicative	X2
3e variable explicative	X4

RESULTATS

Coefficient de corrélation multiple R7771
Carré du coef de corrélation multiple R ² ..	.6039
Valeur du test F de signification de R ² ...	20.3252 avec 3 et 41 ddl
Somme des carrés des écarts SCE	7.5605
Variance résiduelle1890

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 40 ddl
X1	-10.161053	3.368535	3.0164606
X2	.418247	.083591	5.0035116
X4	.040134	.012654	3.1717206
Terme constant	27.039041		

Chimie : Analyse n° 13

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées	44
Nombre de variables utilisées	4
Variable à expliquer	LIAN
2e variable explicative	X3
3e variable explicative	X4

RESULTATS

Coefficient de corrélation multiple R7502
Carré du coef de corrélation multiple R ² ..	.5629
Valeur du test F de signification de R ² ...	17.1681 avec 3 et 41 ddl
Somme des carrés des écarts SCE	8.3431
Variance résiduelle2086

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 40 ddl
X1	-4.499367	3.823973	1.1766216
X3	-.416764	.095776	4.3514256
X4	.055360	.014427	3.8371956
Terme constant	73.832618		

Chimie : Analyse n° 14

=====

REGRESSION MULTIPLE

Nombre d'observations utilisées	44
Nombre de variables utilisées	4
Variable à expliquer	LIAN
2e variable explicative	X3

RESULTATS

Coefficient de corrélation multiple R8285
Carré du coef de corrélation multiple R^2 ..	.6865
Valeur du test F de signification de R^2 ...	29.1939 avec 3 et 41 ddl
Somme des carrés des écarts SCE	5.9839
Variance résiduelle1496

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 40 ddl
X2	.325847	.077449	4.2072716
X3	-.366951	.078170	4.6942686
X4	.070940	.009470	7.4912746

Terme constant 57.290201

Chimie : Analyse n° 15

REGRESSION MULTIPLE

Nombre d'observations utilisées	44
Nombre de variables utilisées	5
Variable à expliquer	LIAN
2e variable explicative	X2
3e variable explicative	X3
4e variable explicative	X4

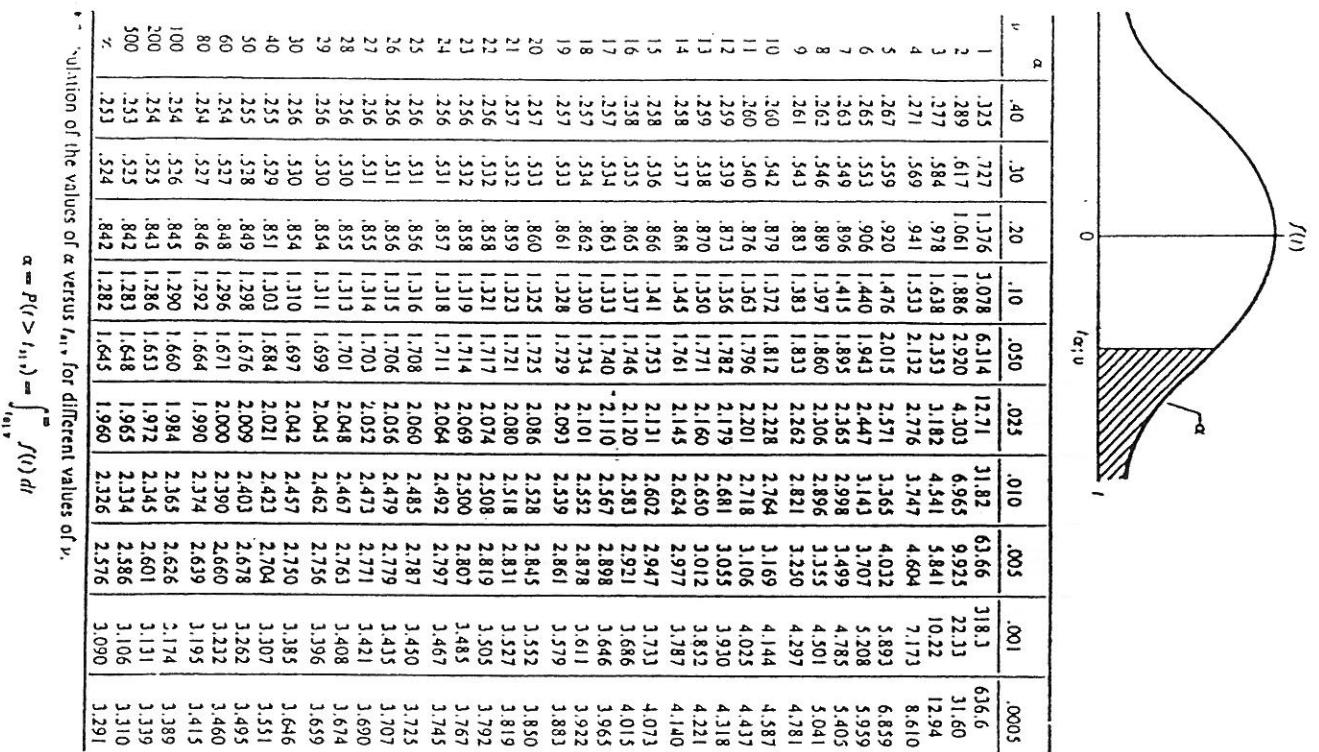
RESULTATS

Coefficient de corrélation multiple R8423
Carré du coef de corrélation multiple R^2 ..	.7095
Valeur du test F de signification de R^2 ...	23.8112 avec 4 et 40 ddl
Somme des carrés des écarts SCE	5.5447
Variance résiduelle1422

Coefficients de régression

Var	Coef	Erreur stand.	Test t de signif. à 39 ddl
X1	-5.564791	3.166209	1.7575566
X2	.335941	.075720	4.4366086
X3	-.310991	.082589	3.7655086
X4	.057680	.011923	4.8378346

Terme constant 56.710798

Table A.3 t distribution [2, 4]•

$$\alpha = P(t > t_{\alpha/2}) = \int_{t_{\alpha/2}}^{\infty} f(t) dt$$

APPENDIX

Table A.4 F distribution [4]Tabulation of the values of $F_{0.10; \nu_1, \nu_2}$ versus ν_1 and ν_2 *

Degrees of freedom for the numerator (ν_1)																		
ν_1	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
1	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	61.2	61.7	62.3	62.7	63.0	63.2	63.3	63.3
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44	9.46	9.47	9.48	9.49	9.49	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18	5.17	5.15	5.14	5.14	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84	3.82	3.80	3.78	3.77	3.76	3.76
5	4.05	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.24	3.21	3.17	3.15	3.13	3.12	3.11	3.10
6	3.78	3.46	3.29	3.18	3.11	3.03	3.01	2.98	2.96	2.94	2.87	2.84	2.80	2.77	2.75	2.73	2.73	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59	2.55	2.52	2.50	2.48	2.48	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42	2.38	2.35	2.32	2.31	2.30	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30	2.25	2.22	2.19	2.17	2.17	2.16
10	3.28	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20	2.16	2.12	2.09	2.07	2.06	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.17	2.12	2.08	2.04	2.00	1.99	1.98	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06	2.01	1.97	1.94	1.92	1.91	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.05	2.01	1.95	1.92	1.88	1.86	1.85	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.01	1.96	1.91	1.87	1.83	1.82	1.80	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92	1.87	1.83	1.79	1.77	1.76	1.76
16	3.03	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.94	1.89	1.84	1.79	1.76	1.74	1.73	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.91	1.86	1.81	1.76	1.73	1.71	1.69	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.89	1.84	1.78	1.74	1.70	1.68	1.67	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.86	1.81	1.76	1.71	1.67	1.65	1.64	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79	1.74	1.69	1.65	1.63	1.62	1.61
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.81	1.76	1.70	1.65	1.61	1.59	1.58	1.57
24	2.93	2.54	2.34	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.78	1.73	1.67	1.62	1.58	1.56	1.54	1.53
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.76	1.71	1.65	1.59	1.55	1.53	1.51	1.50
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.74	1.69	1.63	1.57	1.53	1.50	1.49	1.48
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67	1.61	1.55	1.51	1.48	1.47	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.66	1.61	1.54	1.48	1.43	1.41	1.39	1.38
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76	1.73	1.63	1.57	1.50	1.44	1.39	1.36	1.34	1.33
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.60	1.54	1.48	1.41	1.36	1.33	1.31	1.29
80	2.77	2.37	2.15	2.02	1.92	1.85	1.79	1.75	1.71	1.68	1.57	1.51	1.44	1.38	1.32	1.28	1.26	1.24
100	2.75	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.70	1.66	1.56	1.49	1.42	1.35	1.29	1.26	1.23	1.21
200	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.52	1.46	1.38	1.31	1.24	1.20	1.17	1.14
500	2.72	2.31	2.10	1.96	1.86	1.79	1.73	1.68	1.64	1.61	1.50	1.44	1.36	1.28	1.21	1.16	1.12	1.09
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42	1.34	1.26	1.18	1.13	1.08	1.06

* $F_{\alpha; \nu_1, \nu_2}$ and $F_{\alpha'; \nu_1, \nu_2}$ have been used interchangeably in the text.

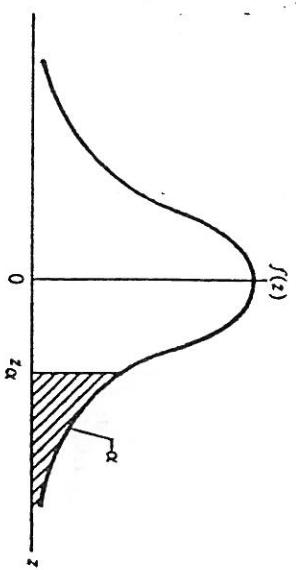


Table A.1 Normal distribution [1]*

z_α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1424	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0750	.0735	.0721	.0708	.0694	.0681

* Tabulation of the values of α versus z_α for the standardized normal curve.

$$\alpha = P(z > z_\alpha) = \int_{z_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

= area under the standardized normal curve from $z = z_\alpha$ to $z = \infty$.

Table A.1 (cont.)

z_α	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
3	.00135	.01988	.03687	.0483	.03337	.03233	.03159	.03108	.04723	.04481
4	.0317	.04207	.04133	.05854	.05341*	.03340	.05211	.05130	.06793	.0479
5	.05287	.06170	.07996	.05759	.07333	.07190	.07107	.08599	.08332	.08182
6	.0987	.09530	.09282	.09149	.07777	.07402	.07206	.07040	.07523	.07260

* .0541 means .00000541.