Kenmerk: TW08/SK/016/kp
Datum: 21 januari 2008

Tentamen van Wiskunde B voor CiT (151217)
Tentamen voor Statistiek voor BIT (153031)
donderdag 31 januari 2008 van 9.00 tot 12.00 uur

Dit tentamen bestaat uit 7 opgaven, 2 tabellen en een formuleblad (2 pagina's).
Vermeld ook je studentnummer op je werk en tentamenbriefje

Opgave 1
Een tentamen bestaat uit meerkeuzevragen, elk met vier mogelijke antwoorden waarvan er slechts 1 goed is. Linda schat dat ze bij elke vraag die gesteld zou kunnen worden een kans van 80% heeft om het goede antwoord te weten. Als ze niet weet wat het goede antwoord is, gaat ze raden en geeft dan met een kans van 25% het goede antwoord.

a. Hoe groot is de kans dat Linda op een willekeurige vraag het goede antwoord geeft?
b. Bepaal de kans dat Linda het goede antwoord op een vraag weet, gegeven dat ze het juiste antwoord gegeven heeft.

Opgave 2
Een producent maakt gloeilampen die met kans 90.5% een levensduur hebben van tenminste 2 (eenheid: duizend uur). We kopen een doos met 24 lampen.

a. Laat Y het aantal gloeilampen (van de 24) zijn dat een levensduur heeft van tenminste 2. Welke verdeling heeft Y? Geef een formule voor $P(Y=20)$. (Uitrekenen hoeft niet.)
b. Stel dat de levensduur van een willekeurige gloeilamp exponentieel is verdeeld met parameter λ. Bereken λ .
c. Benader de kans dat de gemiddelde levensduur van de 24 gloeilampen groter is dan 15. Gebruik $\lambda = 0.0625$ als je opgave *b* niet hebt kunnen oplossen.

Opgave 3
Een stochastische variabele X is verdeeld volgens de kansverdeling gegeven door de volgende kansdichtheid:

$$f(x) = \frac{2}{3} - \frac{2}{9}x \qquad \text{als } 0 < x < 3,$$

en $f(x) = 0$ elders.
Bereken de kans $P(1 < X < 2)$.

Opgave 4
Hoe accuraat zijn de Radon-detectors van een type dat aan huiseigenaren wordt verkocht? Om deze vraag te beantwoorden, plaatsten universitaire onderzoekers 12 detectors in een ruimte waar ze werden blootgesteld aan 105 picocurie Radon per liter (pC/l). De detectoraflezingen waren als volgt. (Gegevens verstrekt door Dina Schellenberg, Purdue University School of Health Sciences.)

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

a. Bereken voor de data het steekproefgemiddelde $\bar{x}$ en de steekproefvariantie s^2.
b. Bereken het 99%-betrouwbaarheidsinterval voor de verwachting van de metingen. Ga uit van een normale verdeling voor de metingen.

Opgave 5
De reputatie (en daarmee de omzet) van veel bedrijven kan zeer nadelig worden beïnvloed door partijen geproduceerde artikelen die een hoog percentage defecte artikelen bevatten. Zo zal een fabrikant van alkaline batterijen er vrij zeker van willen zijn dat minder dan 2% van zijn batterijen defect is. Stel dat er 400 batterijen willekeurig worden gekozen uit een zeer grote partij. Elke batterij wordt getest en er worden 4 defecte gevonden. Is dit voldoende voor de fabrikant om te kunnen concluderen dat de fractie defecte batterijen in de hele partij minder dan 2% is? Gebruik significantieniveau (onbetrouwbaarheidsdrempel) 5% en volg het schema van acht stappen vermeld aan het eind van het tentamen.

Opgave 6
Elektro-encefalogrammen tonen de elektrische activiteit in de hersenen. Van de verschillende te onderscheiden golven in de hersenen zijn de zogenaamde alfagolven dominant. Deze golven hebben een frequentie van 8 tot 13 cycli per seconde. Het volgende experiment is uitgevoerd om te onderzoeken of opsluiting in een isoleercel in een gevangenis het patroon van de alfagolven beïnvloedt.
Twintig gevangenen van een Canadese gevangenis werden via een lotingsysteem lukraak verdeeld in twee groepen van 10. De leden van groep 1 moesten (afzonderlijk) in isoleercellen verblijven. De leden van de andere groep mochten in de eigen cel blijven. Na zeven dagen werd de frequentie van de alfagolven gemeten voor alle twintig gevangenen. De resultaten zijn als volgt.

in isoleercel	9.6	10.4	9.7	10.3	9.2	9.3	9.9	9.5	9.0	10.9
in eigen cel	10.7	10.7	10.4	10.9	10.5	10.3	9.6	11.1	11.2	10.4

De steekproefgemiddelden zijn respectievelijk 9.78 (isoleercel) en 10.58 (eigen cel). De steekproefstandaardafwijkingen zijn respectievelijk 0.598 (isoleercel) en 0.459 (eigen cel).
Bereken het 99%-betrouwbaarheidsinterval voor het verschil in verwachtingen tussen (de verdelingen van) de twee steekproeven. Ga uit van normale verdelingen voor de twee steekproeven en gelijke varianties.

Opgave 7
We bestuderen hoe de weerstand van rubber tegen afschuren beïnvloed wordt door de hardheid (x_1) en treksterkte (x_2) van rubber. Dertig monsters zijn getest op hardheid (in graden Shore, hoe hoger het getal des te harder is het rubber) en treksterkte (gemeten in kg/cm^2). Daarna is elk monster rubber een tijd lang afgeschuurd (alle monsters op dezelfde wijze) en is het afgeschuurde gewicht aan rubber per uur (y) bepaald (eenheid: g/uur). De data zijn als volgt:

y	x_1	x_2	y	x_1	x_2
372	45	162	196	68	173
206	55	233	128	75	188
175	61	232	97	83	161
154	66	231	64	88	119
136	71	231	249	59	161
112	71	237	219	71	151
55	81	224	186	80	165
45	86	219	155	82	151
221	53	203	114	89	128
166	60	189	341	51	161
164	64	210	340	59	146
113	68	210	283	65	148
82	79	196	267	74	144
32	81	180	215	81	134
228	56	200	148	86	127

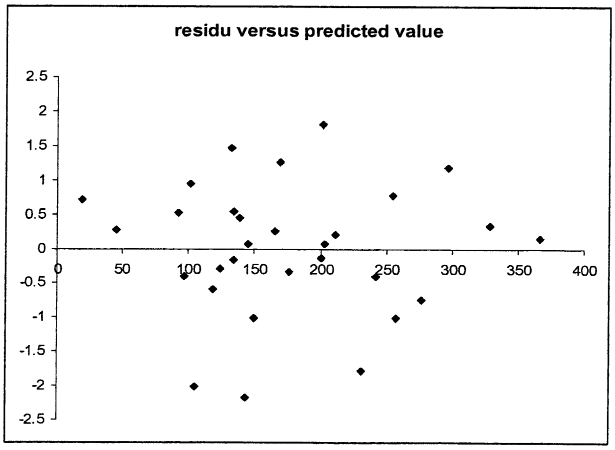
We gaan uit van het model van meervoudige lineaire regressie met hardheid (x_1) en treksterkte (x_2) als verklarende variabelen. Met het pakket GENSTAT hebben we de volgende output gevonden:

Summary of analysis/Analysis of variance

	df	SS	MS	F
Regression	2	189062	94531	71.10
Residual	27	35950	1331	
Total	29	225011		

Estimates of regression coefficients

	estimate	s.e.	t
constant	885.2	61.8	14.33
hardheid	-6.571	0.583	-11.27
treksterkte	-1.374	0.194	-7.07



a. Op de vorige bladzijde vind je een plot waarin de (gestandaardiseerde) residuen (verticaal) zijn uitgezet tegen de waarden van de 'predicted values' $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$. Geef commentaar. Past het (meervoudige lineaire) regressiemodel goed bij de data?

b. Bereken R_a^2 (adjusted R square).

c. Je kunt je afvragen of het zin heeft treksterkte (x_2) te handhaven in het model als je al hardheid (x_1) als verklarende variabele hebt. Voer een toets uit om na te gaan of treksterkte (x_2), in toevoeging op hardheid (x_1), van invloed is op de afhankelijke variabele y. Volg het schema van 8 stappen vermeld aan het eind van het tentamen en gebruik significantieniveau (onbetrouwbaarheidsdrempel) 1 %.

Normering:

1a	1b	2a	2b	2c	3	4a	4b	5	6	7a	7b	7c
2	3	2	2	4	2	3	3	4	3	2	2	4

Totaal: 36

Schema van acht stappen.

1. Formuleer het kansmodel.
2. Formuleer nulhypothese H_0 en alternatieve hypothese H_a in termen van de parameters van het kansmodel.
3. Formuleer een geschikte toetsingsgrootheid in termen van de voorkomende s.v.-en.
4. Geef de kansverdeling van de toetsingsgrootheid onder (het randpunt van) H_0
5. Bereken of geef de waarde van de toetsingsgrootheid.
6. Bepaal de kritieke waarde(n) en geef het kritieke gebied.

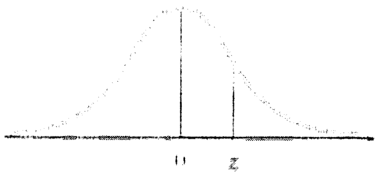
of

6^*. Bereken de overschrijdingskans.

7. Formuleer de conclusie omtrent het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheid(sdrempel).
8. Vermeld de conclusie in "gewone woorden".

Bijlagen:
Formuleblad voor de vakken 151217 en 153031 (2 pagina's)
Tabel standaardnormale verdeling
Tabel t-verdeling

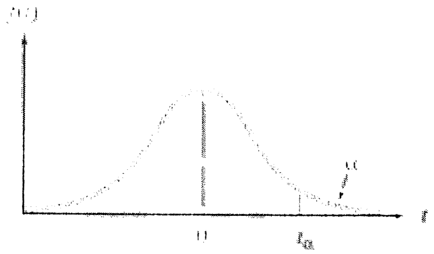
Tabel IV Oppervlakten bij de standaardnormale verdeling



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Bron: Verkort overgenomen uit tabel II van A. Hald, *Statistical Tables and Formulas* (New York: Wiley), 1952. Overgenomen met toestemming van A. Hald.

Tabel VI Kritieke waarden van t



Aantal vrijheidsgraden	$t_{0,100}$	$t_{0,050}$	$t_{0,025}$	$t_{0,010}$	$t_{0,005}$	$t_{0,001}$	$t_{0,0005}$
1	3,078	6,314	12,706	31,821	63,657	318,31	636,62
2	1,886	2,920	4,303	6,965	9,925	22,326	31,598
3	1,638	2,353	3,182	4,541	5,841	10,213	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	1,303	1,684	2,021	2,423	2,704	3,307	3,551
60	1,296	1,671	2,000	2,390	2,660	3,232	3,460
120	1,289	1,658	1,980	2,358	2,617	3,160	3,373
∞	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Bron: Deze tabel is overgenomen met toestemming van de Trustees van Biometrika uit E.S. Pearson en H.O. Hartley (eds.) *The Biometrika Tables for Statisticians*, Vol. I, 3d ed., Biometrika, 1966.

Formuleblad voor de vakken 151217 en 153031 deel 1/2

Verdeling	verwachting	variantie
Binomiaal	np	$np(1-p)=npq$
Poisson	λ	λ
exponentieel	$1/\lambda$	$1/\lambda^2$

Normale benadering van binomiale verdeling: toepassen als $np \geq 5$ en $n(1-p) \geq 5$.

$E(X+Y)=E(X)+E(Y)$

$E(aX+b)=aE(X)+b$

$\mathrm{var}(X+Y)=\mathrm{var}(X)+\mathrm{var}(Y)$ als X en Y onafhankelijk zijn

$\mathrm{var}(aX+b)=a^2\,\mathrm{var}(X)$

$\mathrm{var}(X-Y)=\mathrm{var}(X)+\mathrm{var}(Y)$ als X en Y onafhankelijk zijn

1 steekproef

Betrouwbaarheidsintervallen (BI's): $\overline{X} \pm t_{\alpha/2}\dfrac{S}{\sqrt{n}}$ en $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

Toetsingsgrootheden: $T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$ en $Z=\dfrac{\hat{p}-p_0}{\sigma_{\hat{p}}}=\dfrac{\hat{p}-p_0}{\sqrt{p_0q_0/n}}$

2 steekproeven

BI's: $\overline{X}_1-\overline{X}_2 \pm t_{\alpha/2}\sqrt{S_p^2(\frac{1}{n_1}+\frac{1}{n_2})}$ en $\hat{p}_1-\hat{p}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1}+\dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

Toetsingsgrootheden: $T=\dfrac{\overline{X}_1-\overline{X}_2-D_0}{\sqrt{S_p^2(\frac{1}{n_1}+\frac{1}{n_2})}}$ en $Z=\dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}\hat{q}(\frac{1}{n_1}+\frac{1}{n_2})}}$

Met $S_p^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ en $\hat{p}=\dfrac{X_1+X_2}{n_1+n_2}$

$H_0:\dfrac{\sigma_1^2}{\sigma_2^2}=1$ tegen $H_a:\dfrac{\sigma_1^2}{\sigma_2^2}\neq 1$, Toetsingsgrootheid: $F=\dfrac{S_1^2}{S_2^2}$

H_0 verwerpen als (grootste steekproefvariantie)/(kleinste steekproefvariantie) $\geq F_{\alpha/2}$

Formuleblad voor de vakken 151217 en 153031 deel 2/2

Enkelvoudige lineaire regressie

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \qquad \hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} \qquad r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

$$SS_{xx} = \sum_i (x_i - \bar{x})^2 \qquad SS_{yy} = \sum_i (y_i - \bar{y})^2 \qquad SS_{xy} = \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

Toetsingsgrootheid: $T = \dfrac{\hat{\beta}_1}{S/\sqrt{SS_{xx}}}$

Betrouwbaarheidsinterval voor β_1: $\hat{\beta}_1 \pm t_{\alpha/2}\dfrac{S}{\sqrt{SS_{xx}}}$

Betrouwbaarheidsinterval voor $E(y)$ voor $x = x_p$: $\hat{y} \pm t_{\alpha/2} S \sqrt{\dfrac{1}{n} + \dfrac{(x_p - \bar{x})^2}{SS_{xx}}}$

Voorspellingsinterval voor nieuwe y voor $x = x_p$: $\hat{y} \pm t_{\alpha/2} S \sqrt{1 + \dfrac{1}{n} + \dfrac{(x_p - \bar{x})^2}{SS_{xx}}}$

Meervoudige lineaire regressie

$$S^2 = \frac{SSE}{n-(k+1)} = \frac{\sum_i (y_i - \hat{y}_i)^2}{n-(k+1)}$$

$$R^2 = 1 - \frac{SSE}{SS_{yy}} \qquad R_a^2 = 1 - \left(\frac{n-1}{n-(k+1)}\right)\frac{SSE}{SS_{yy}}$$

Betrouwbaarheidsinterval voor β_i: $\hat{\beta}_i \pm t_{\alpha/2} S_{\hat{\beta}_i}$

Toetsingsgrootheden: $T = \dfrac{\hat{\beta}_i}{S_{\hat{\beta}_i}}$ en $F = \dfrac{(SS_{yy} - SSE)/k}{SSE/(n-(k+1))} = \dfrac{R^2/k}{(1-R^2)/(n-(k+1))}$