

Tentamen van Wiskunde B voor CiT (151217)
Tentamen van Statistiek voor BIT (153031)
donderdag 29 januari 2009 van 9:00 tot 12:00 uur

*Dit tentamen bestaat uit 8 opgaven, 2 tabellen en een formuleblad (2 pagina's).
Vermeld ook je studentnummer op je werk en tentamenbriefje.*

Opgave 1

Een firma brengt een offerte uit om een groot project binnen te halen. Het management van de firma schat in dat de kans om het project te krijgen 60% is. Na indiening van het bod kan de beoordelende instantie die het project toewijst om extra informatie vragen. Uit het verleden is bekend dat bij 75% van de gehonoreerde offertes om extra informatie is gevraagd en dat bij de niet gehonoreerde offertes in 40% van de gevallen om extra informatie is verzocht. Er wordt om extra informatie gevraagd. Bereken met behulp van dit gegeven de kans dat de offerte door de firma voor het project gehonoreerd wordt. Definieer daartoe eerst een aantal relevante gebeurtenissen en geef de gegeven kansen weer in termen van (voorwaardelijke) kansen op die gebeurtenissen.

Opgave 2

De stochastische variabele X heeft de volgende kansverdeling.

x	0	1	2	3	4
$P(X = x)$	0.34	0.40	0.20	0.04	0.02

- a. Bereken $\mu = E(X)$.
- b. Bereken $\sigma^2 = \text{var}(X)$.
- c. Bereken $P(X \geq 3 | X \geq 2)$.

Opgave 3

Een machine die stansstukken maakt voor automotoren functioneert niet goed en levert 10% defecte producten. De volgorde waarin defecte en gave stansstukken door de machine worden geproduceerd is willekeurig. Bepaal de kans dat van de tien volgende stansstukken er (precies) vier defect zijn.

Opgave 4

De gipskamers van ziekenhuizen hadden het in de afgelopen vorstperiode soms erg druk. Veronderstel dat de tijdsduur (eenheid: uur) van de behandeling van een botbreuk een exponentiële verdeling heeft met parameter $\lambda = 3$. Voor de tijdsduur van een behandeling geldt dus $\mu = 1/3$ en $\sigma = 1/3$. Stel dat een gipskamer op een dag 30 mensen met een botbreuk behandelt. Bereken/benader met behulp van de Centrale Limietstelling de kans dat de totale behandelingsduur van de 30 botbreuken meer dan 12 uur bedraagt.

Opgave 5

Een verzekeringsmaatschappij kan zich voor de reparatie van autoschades wenden tot twee garages. Om te onderzoeken welke van de twee garages het goedkoopst is laat men van 7 willekeurige auto's de schade taxeren door zowel garage 1 als garage 2. Onderstaande tabel geeft voor elke auto en voor elk van de twee garages de getaxeerde bedragen (in honderden euro's).

auto	1	2	3	4	5	6	7
garage 1	15.2	20.4	19.0	6.2	8.0	12.6	10.6
garage 2	14.7	18.3	16.8	6.7	7.6	11.6	9.8

Onderzoek met een geschikte statistische toets of er verschil is in verwachte taxatie bedragen tussen de twee garages. Neem als significantieniveau $\alpha = 0.05$ en volg het schema van 8 stappen vermeld aan het eind van het tentamen.

Opgave 6

In het kader van de studeerbaarheid van studieprogramma's wordt onderzocht hoeveel tijd studenten op een werkdag aan de studie besteden. Veronderstel dat studietijden normaal verdeeld en onafhankelijk zijn. Bij een onderzoek onder 100 studenten bedraagt het steekproefgemiddelde 408.73 minuten en de steekproefstandaardafwijking 57.39 minuten.

- Bereken het 95%-betrouwbaarheidsinterval voor de verwachting van de studieduur.
- Voor 95% van de studenten zal de studietijd in het interval van onderdeel *a* liggen (indien dit correct berekend is). Is deze bewering of waar of niet waar? Motiveer kort je antwoord.

Opgave 7

Bestaat er tussen mannelijke en vrouwelijke promovendi een verschil in prestaties? Een grote universiteit classificeerde alle promovendi die in een bepaald jaar begonnen met een promotieonderzoek volgens de 6 jaar later bereikte status. Van de 229 vrouwen van het onderzoek waren er 98 gepromoveerd na 6 jaar. Van de 795 mannen van het onderzoek waren er 423 gepromoveerd na 6 jaar. Voer een toets uit om na te gaan of er tussen mannelijke en vrouwelijke promovendi een verschil bestaat in het percentage promoties na 6 jaar. Neem 5% als significantieniveau en volg het schema van acht stappen vermeld aan het eind van dit tentamen.

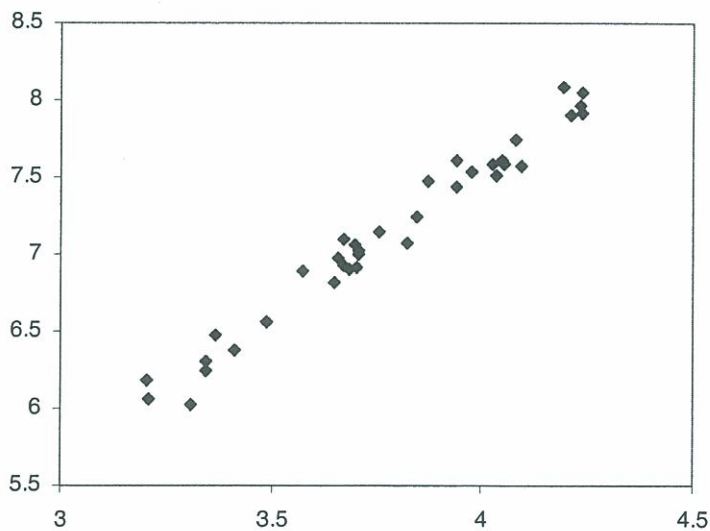
Opgave 8

Hardheid (volgens de definitie van Janka) is een belangrijke structurele eigenschap van (timmer)hout en moeilijk te meten. Hardheid van hout is echter gerelateerd aan dichtheid van hout en de dichtheid is relatief gemakkelijk te meten. Voor Australisch hardhout (Eucalyptus) willen we de relatie tussen hardheid en dichtheid bepalen teneinde de hardheid op basis van de dichtheid te kunnen voorspellen/schatten. We bestuderen de metingen van hardheid (volgens Janka) en dichtheid van 36 monsters Australisch hardhout. De metingen zijn als volgt:

dichtheid	hardheid	dichtheid	hardheid	dichtheid	Hardheid
24.7	484	39.4	1210	53.4	1880
24.8	427	39.9	989	56.0	1980
27.3	413	40.3	1160	56.5	1820
28.4	517	40.6	1010	57.3	2020
28.4	549	40.7	1100	57.6	1980
29.0	648	40.7	1130	59.2	2310
30.3	587	42.9	1270	59.8	1940
32.7	704	45.8	1180	66.0	3260
35.6	979	46.9	1400	67.4	2700
38.5	914	48.2	1760	68.8	2890
38.8	1070	51.5	1710	69.1	2740
39.3	1020	51.5	2010	69.1	3140

We passen (enkelvoudige) lineaire regressie toe met $y = \ln(\text{hardheid})$, de natuurlijke logaritme van hardheid, als afhankelijke variabele en met $x = \ln(\text{dichtheid})$ als verklarende variabele. De variabele y is uitgezet tegen x in Figuur 1.

Figuur 1: y versus x



Computer output is als volgt:

ANOVA

	Sum of Squares	df	Mean Square	F
Regression	11.7678	1	11.767781	1231.36
Residual	0.3249	34	0.009557	
Total	12.0927	35	0.345506	

Coefficients

	B	Std. Error	T
Constant	0.015	0.204	0.08
x	1.8847	0.0537	35.09

- Geef het kansmodel van de enkelvoudige lineaire regressiemodel voor deze situatie.
- Bereken de correlatiecoëfficiënt r .
- geef de (gebruikelijke) schatting van σ^2 .
- Voor een nieuw monster Australisch hardhout (Eucalyptus) wordt de waarde 4.0 voor $x = \ln(\text{dichtheid})$ gevonden. Bereken een interval zodanig dat de bijbehorende waarde van de hardheid van dit monster met kans 90% in dit interval zal vallen. Maak voor dit interval gebruik van de computer output en de volgende gegevens:

Steekproefgemiddelde van x : 3.77811

Steekproefvariantie van x : 0.09465

Normering.

1	2a	2b	2c	3	4	5	6a	6b	7	8a	8b	8c	8d
3	2	2	2	2	4	6	2	2	4	2	2	1	4

Totaal: 38 punten

Schema van acht stappen.

- Formuleer het kansmodel.
- Formuleer nulhypothese H_0 en alternatieve hypothese H_1 in termen van de parameters van het kansmodel.
- Formuleer een geschikte toetsingsgrootheid in termen van de voorkomende s.v.-en.
- Geef de kansverdeling van de toetsingsgrootheid onder (het randpunt van) H_0
- Bereken of geef de waarde van de toetsingsgrootheid.
- Bepaal de kritieke waarde(n) en geef het kritieke gebied.
of
- *. Bereken de overschrijdingskans.
- Formuleer de conclusie omtrent het al dan niet verwerpen van H_0 bij de gegeven onbetrouwbaarheid(sdrempel).
- Vermeld de conclusie in "gewone woorden".

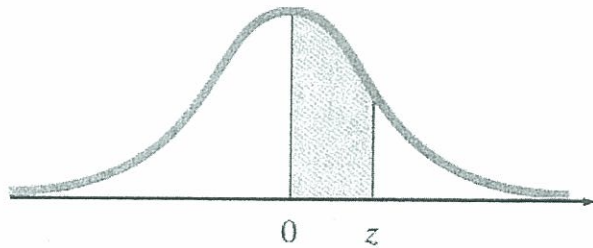
Bijlagen:

Formuleblad voor de vakken 151217 en 153031 (2 pagina's)

Tabel standaardnormale verdeling

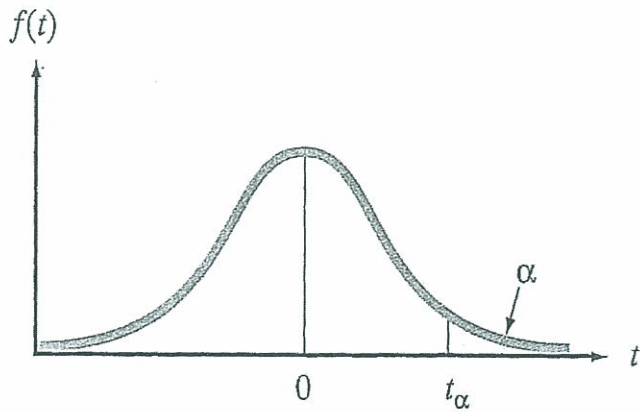
Tabel t -verdeling

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

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	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	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	3.768
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

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 $n(1-p) \geq 5$.

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

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

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

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

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

1 steekproef

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

$$\text{Toetsingsgrootheden: } T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad \text{en} \quad Z = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}} = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

2 steekproeven

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

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

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

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

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

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

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

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

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

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