Tentamen Kansrekening en Statistiek voor Technische Aardwetenschappen, 28 juni 2002

Het tentamen bestaat *alleen* uit *open vragen*. Alle antwoorden moeten zijn voorzien van een (korte) argumentatie, dus laat bij elke vraag zien hoe je aan het antwoord bent gekomen. Een vraag waarbij alleen het antwoord is vermeld, wordt niet goed gerekend. Voor meer duidelijkheid omcirkel de eindelijke antwoorden. Niet alle vragen tellen even zwaar. Naast elke vraag staat hoeveel punten deze vraag waard is. Als je een deel van een vraag goed hebt, krijg je deel van de punten. Het maximale mogelijke aantal punten is 36, echter je hoeft maar 30 punten scoren (d.w.z. wie 30 punten scoort, krijgt cijfer 10). Dus je hoeft niet alle vragen te doen. Je mag zelf bepalen welke vragen je wel of niet doet.

Gedurende het tentamen mag een zelfgemaakt uittreksel van hooguit twee A4 gebruikt worden en een rekenmachine.

1. *(2 punten)* Iemand doet twee worpen met een zuivere dobbelsteen.
   A is de gebeurtenis "de som van de ogen is gelijk aan 4";
   B is de gebeurtenis "bij ten minste één van de worpen komt er 3 boven".

   a) Bereken de kans $P(A|B)$.

   b) Zijn A en B onafhankelijk? Motiveer je aantwoord.

2. *(2 punten)* Men kiest net zolang zonder teruglegging een getal uit de verzameling $\{1, 2, 3\}$ totdat het getal 3 getrokken wordt. De stochast X is het aantal benodigde trekkingen.

   a) Geef de kansverdeling van X.

   b) Bereken EX en Var$X$.

3. *(2 punten)* Stochast X heeft een uniforme verdeling op het interval [0,3], en stochast Y heeft een uniforme verdeling op het interval [1,4]. X en Y zijn onafhankelijk. Bereken de kans $P(Y \leq X)$. (Hint: gebruik meetkundige interpretatie van de kans.)

4. *(3 punten)* X en Y zijn onafhankelijke discrete stochasten met dezelfde verdeling $\begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$.

   Wij definiëren nieuwe stochasten $U = min(X, Y)$ en $V = max(X, Y)$.

   a) Geef (in de form van een tabel) de simultane dichtheid van U en V.

   b) Bereken EU en EV.

   c) Bereken de covariantie tussen U en V. Zijn U en V onafhankelijk, ongecorreleerd of beide?

5. *(3 punten)* De continue stochast X heeft de volgende verdelingsfunctie:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x^2, & 0 < x \leq 1, \\ 1, & x > 0. \end{cases}$$

   Bereken: de dichtheid $f(x)$, EX, $E(X^3)$, Var$(2X)$ en de verdelingsfunctie van de stochast $Y = X^3$.

6. *(3 punten)* De gezamenlijke dichtheid van de stochasten X en Y wordt gegeven door

$$f(x, y) = \begin{cases} x + y, & 0 < x < 1, \quad 0 < y < 1 \\ 0, & \text{anders} \end{cases}$$

   a) Wat zijn de marginale dichtheden van X en van Y?

   b) Zijn X en Y onafhankelijk? (motiveer jouw antwoord!)

   c) Bereken EX, EY en $E(XY)$.

7. *(2 punten)* Stel X en Y zijn onafhankelijke stochasten, beide binomiaal verdeeld met parameters n en p, en Z is gedefinieerd door $Z = X + Y$. Wat is de verdeling van Z?

8. *(2 punten)* Een machine vervaardigt onderdelen waarvan de lengte normaal verdeeld is met gemiddelde 10 cm en standaardafwijking $\sigma = 2$ cm. Een onderdeel wordt afgekeurd als de lengte meer dan 1.5 cm van het gemiddelde afwijkt. Welk percentage van de gefabriceerde onderdelen zal op lange duur worden afgekeurd?

   N.B. Op de volgende bladzijde staat een tabel van de normale verdeling afgedrukt.

9. *(2 punten)* Iemand heeft de beschikking over een random generator die realisaties $w_1, w_2, \ldots$ uit een uniforme verdeling op $[0,1]$ produceert. Hoe kan zij hiermee realisaties simuleren van een stochast X met de volgende continue dichtheid: $f(x) = \frac{1}{2\sqrt{x}}$, voor $x \in (0,1]$, en $f(x) = 0$ voor $x \notin (0,1]$?

10. *(4 punten)* Laat gegeven zijn dat de stochasten $X_1, X_2, \ldots, X_{100}$ onderling onafhankelijk zijn, en allemaal dezelfde verwachting $m = 75$ en dezelfde variantie $\sigma^2 = 225$ hebben. Men berekent het gemiddelde van deze stochasten: $\bar{X}_{100}$.

    a) Geef een afschatting van de kans $P(|\bar{X}_{100} - m| > 3)$ met behulp van de Chebyshev ongelijkheid.

    b) Geef een benadering van dezelfde kans met behulp van de Centrale Limietstelling.

11. *(3 punten)* De dataset $x_1, x_2, \ldots, x_n$ is afkomstig van n onafhankelijke stochasten $X_1, \ldots, X_n$ met een zogenoemde gamma$(2,\theta)$ verdeling, met onbekende θ. Dat wil zeggen

$$f(x) = \begin{cases} \theta^2 x\, e^{-\theta x}, & \text{als } x \geq 0 \\ 0, & \text{als } x < 0 \end{cases}$$

    Bepaal de maximum likelihood schatter voor θ.

12. *(4 punten)* De stochasten $X_1, X_2, \ldots, X_n$ zijn onafhankelijk en hebben dezelfde verdeling. Voor een vaste t, de stochast $\hat{F}_n(t)$ is de empirische verdelingsfunctie, d.w.z.

$$\hat{F}_n(t) = \frac{\text{aantal } X_i \leq t}{n}.$$

    a) Bereken $E(\hat{F}_n(t))$. Hint: gebruik indicator stochasten: $Y_i = 1$ als $X_i \leq t$ en $Y_i = 0$ als $X_i > t$.

    b) Bereken de variantie van $\hat{F}_n(t)$.

13. *(4 punten)* Een bepaalde soort onderdeel hoort 35 mm lang te zijn. Metingen van 20 willekeurig gekozen onderdelen geven de volgende resultaten aan:

Lengte:	34.8	34.9	35.0	35.1	35.3	(in mm)
Aantal:	2	3	4	6	5	

    (Dat betekent bijv. dat van 20 onderdelen 2 zijn 34.8 mm lang, 3 zijn 34.9 mm lang enz.) Aannemende dat deze metingen afkomstig zijn uit een normale $N(\mu, \sigma^2)$ verdeling, waarbij μ en σ onbekend zijn, willen wij toetsen

$$H_0 : \mu = 35 \quad \text{tegen} \quad H_1 : \mu \neq 35$$

    bij significantieniveau $\alpha = 0.05$.

    a) Wat is de toetsingsgrootheid T? Bepaal zijn waarde op grond van de data en geef de kritieke waarden en het kritieke gebied van deze toets.

    b) Wordt de nulhypothese verworpen of niet? Waarom?

    c) Geef het 95% betrouwbaarheidsinterval voor μ op grond van deze metingen.

z	0	1	2	3	4	5	6	7	8	9
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	1423	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	0749	0735	0721	0708	0694	0681
1.5	0668	0655	0643	0630	0618	0606	0594	0582	0571	0559
1.6	0548	0537	0526	0516	0505	0495	0485	0475	0465	0455
1.7	0446	0436	0427	0418	0409	0401	0392	0384	0375	0367
1.8	0359	0351	0344	0336	0329	0322	0314	0307	0301	0294
1.9	0287	0281	0274	0268	0262	0256	0250	0244	0239	0233
2.0	0228	0222	0217	0212	0207	0202	0197	0192	0188	0183
2.1	0179	0174	0170	0166	0162	0158	0154	0150	0146	0143
2.2	0139	0136	0132	0129	0125	0122	0119	0116	0113	0110
2.3	0107	0104	0102	0099	0096	0094	0091	0089	0087	0084
2.4	0082	0080	0078	0075	0073	0071	0069	0068	0066	0064
2.5	0062	0060	0059	0057	0055	0054	0052	0051	0049	0048
2.6	0047	0045	0044	0043	0041	0040	0039	0038	0037	0036
2.7	0035	0034	0033	0032	0031	0030	0029	0028	0027	0026
2.8	0026	0025	0024	0023	0023	0022	0021	0021	0020	0019
2.9	0019	0018	0018	0017	0016	0016	0015	0015	0014	0014
3.0	0013	0013	0013	0012	0012	0011	0011	0011	0010	0010
3.1	0010	0009	0009	0009	0008	0008	0008	0008	0007	0007
3.2	0007	0007	0006	0006	0006	0006	0006	0005	0005	0005
3.3	0005	0005	0005	0004	0004	0004	0004	0004	0004	0003
3.4	0003	0003	0003	0003	0003	0003	0003	0003	0003	0002

Tabel 1: *Rechteroverschrijdingskansen* $\mathbf{P}(Z \geq z)$ *van de standaard normale verdeling.*

$k \backslash \alpha$	0.1	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
∞	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Tabel 2: *Kritieke waarden $t_{k,\alpha}$ van de Student verdeling*

Uitwerkingen van tentamen
Statistiek wi2068Ta op 28.07.02

Opgave 1 Totaal 36 uitkomsten (11, 12, 13, …, 65, 66)

Geb. A: $\{13, 31, 22\}$, dus $P(A) = \frac{3}{36} = \frac{1}{12}$

Geb. B: $\{13, 31, 23, 32, 33, 43, 34, 53, 35, 63, 36\}$, dus $P(B) = \frac{11}{36}$

Geb. $A \cap B$ (A en B): $\{31, 13\}$, dus $P(A \cap B) = \frac{2}{36}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/36}{11/36} = \frac{2}{11}$$

) onafhankelijk als $P(A \cap B) = P(A) \cdot P(B)$.
Wij hebben: $P(A) \cdot P(B) = \frac{1}{12} \cdot \frac{11}{36} \neq \frac{2}{36} = P(A \cap B)$

dus A en B zijn afhankelijk.

Opgave 2 Verz. $\{1, 2, 3\}$, trekkingen zonder teruglegging, dus kunnen maximaal 3 keer trekken. Zij $X = \#$ trekkingen tot 3.

$P(X=1) = \frac{1}{3}$ (3 bij de eerste trek)

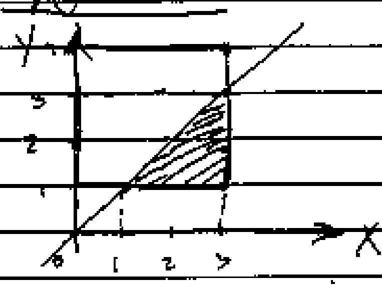
$P(X=2) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$ (1 of 2 bij eerste, 3 bij tweede tr.)

$P(X=3) = \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{3}$ (1 of 2 bij eerste, geen 3 bij 2e)

$\mathbb{E}X = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 3 = 2$

$\mathrm{Var}X = \mathbb{E}(X - \mathbb{E}X)^2 = (-1)^2 \cdot \frac{1}{3} + 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} = \frac{2}{3}$

Opgave 3 $X \sim Un[0,3]$, $Y \sim Un[1,4]$. $P(Y \leq X)$ - ?



Dus paar (X, Y) neemt waarden in de vierkant $(0,1), (0,4), (3,4), (3,1)$

$Y \leq X$ als paar (X, Y) onder de grafiek van $y = x$ ligt, dus in de gestippelde triehoek. De kans $P(Y \leq X)$ is dan gelijk aan:

$$P(Y \leq X) = \frac{\text{oppervlak van } \blacktriangle}{\text{Totale oppervlak}} = \frac{2}{9}$$

Opgave 4 $X, Y \sim \begin{pmatrix} 0 & 1 \\ 1/3 & 2/3 \end{pmatrix}$ $U = \min(X, Y)$, $V = \max(X, Y)$.

a) Gezamenlijke verdeling van U en V is berekend volgens bijv.:
$P(U=0, V=0) = P(X=0, Y=0) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$
$P(U=0, V=1) = P(X=0, Y=1) + P(X=1, Y=0) = \frac{4}{9}$ en z.

Dus wij krijgen het volgende tabel:

u \ v	0	1	
0	1/9	4/9	5/9
1	0	4/9	4/9
	1/9	8/9	1

waar in de laatste rij en colom staan marginale verdelingen van U en V.
Dus verder berekenen wij:

b) $\mathbb{E}U = 0 \cdot \frac{5}{9} + 1 \cdot \frac{4}{9} = \frac{4}{9}$, $\mathbb{E}V = 0 \cdot \frac{1}{9} + 1 \cdot \frac{8}{9} = \frac{8}{9}$

c) $\mathrm{Cov}(UV) = \mathbb{E}(U \cdot V) - \mathbb{E}U \cdot \mathbb{E}V = \frac{4}{9} - \frac{4}{9} \cdot \frac{8}{9} = \frac{4}{81}$
U en V zijn gecorreleerd want $\mathrm{cov}(UV) \neq 0$ en dus ook afhankelijk.

Opgave 5

$F(x) = \begin{cases} 0, & x \le 0 \\ x^2, & x \in (0,1] \\ 1, & x > 1 \end{cases}$

$f(x) = F'(x) = \begin{cases} 2x, & x \in (0,1) \\ 0, & \text{anders} \end{cases}$

$\mathbb{E}X = \int_0^1 x \cdot f(x)\,dx = \int_0^1 2x^2\,dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3}$

$\mathbb{E}X^3 = \int_0^1 x^3 f(x)\,dx = \int_0^1 2x^4\,dx = \frac{2}{5}x^5\Big|_0^1 = \frac{2}{5}$

$\mathrm{Var}\,2X = 4\,\mathrm{Var}X$; $\quad \mathrm{Var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2$

$\mathbb{E}X^2 = \int_0^1 x^2 f(x)\,dx = \int_0^1 2x^3\,dx = \frac{2}{4}x^4\Big|_0^1 = \frac{1}{2}$

Dus $\mathrm{Var}\,2X = 4 \cdot \left(\frac{1}{2} - \frac{4}{9}\right) = 4 \cdot \frac{1}{18} = \frac{4}{18} = \frac{2}{9}$

$Y = X^3$: $F_Y(y) = P(Y \le y) = P(X^3 \le y) =$

$= P(X \le y^{1/3}) = F_X(y^{1/3}) = (y^{1/3})^2 = y^{2/3}$ als $y \in (0,1]$

(en 0 als $y < 0$, 1 als $y > 1$)

Opgave 6.

$f(x,y) = \begin{cases} x+y, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{elders} \end{cases}$

a) $f_X(x) = \int_0^1 (x+y)\,dy = \int_0^1 x\,dy + \int_0^1 y\,dy = x + \frac{1}{2}$

$f_Y(y) = y + \frac{1}{2}$ (hetzelfde berekening).

b) Als X en Y onafhankelijk zijn, dan $f(x,y) = f_X(x) \cdot f_Y(y)$

wij hebben: $f(x,y) = x+y \ne (x+\frac{1}{2}) \cdot (y+\frac{1}{2}) = f_X(x) \cdot f_Y(y)$

Dus X en Y zijn <u>afhankelijk</u>.

c) $\mathbb{E}X = \int_0^1 x f_X(x)\,dx = \int_0^1 x(x+\frac{1}{2})\,dx = \frac{7}{12}$

$\mathbb{E}Y = \frac{7}{12}$ (hetzelfde berekening)

$\mathbb{E}(XY) = \int_0^1\int_0^1 xy\,(x+y)\,dx\,dy = \int_0^1\int_0^1 (x^2y + y^2x)\,dx\,dy =$

$=$ (berekening) $= \frac{1}{3}$

Opgave 7

$X, Y \sim \mathrm{Bin}(n,p)$; $Z = X+Y$

Dat X en Y binomiaal verdeeld zijn, betekent dat ze allebei het aantal successen zijn in twee onafhankelijke rijen van n experimenten, iedere met succeskans p. Dus na een beetje nadenken zien wij dat $X+Y$ het aantal successen zijn in een rij van $2n$ experimenten met succeskans p, dus $\sim \mathrm{Bin}(2n,p)$

Opgave 8

Lengte $L \sim N(10,2)$

Gevraagd: $P(|L-10| > 1.5)$ -?

$P(|L-10| > 1.5) = P(L-10 > 1.5) + P(10-L > 1.5) =$

$= 1 - P(8.5 \le L \le 11.5) = 1 - P(-1.5 \le L-10 \le 1.5) =$

$= 1 - P\left(-\frac{1.5}{2} \le \underbrace{\frac{L-10}{2}}_{Z\sim N(0,1)} \le \frac{1.5}{2}\right) = 1 - P(-0.75 \le Z \le 0.75)$

waar $Z \sim N(0,1)$.

Uit tabel: $1 - P(-0.75 \le Z \le 0.75) = 2 \cdot 0.2266 = 0.4532$

Opgave 9 $\omega_1, \omega_2 \ldots \sim Un[0,1]$

Wij willen genereren random getallen uit verdeling met dichtheid

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & x \in (0,1) \\ 0, & \text{elders} \end{cases}$$

Berekenen verdelingsfunctie:

$$F(x) = \int_0^x f(t)\,dt = \int_0^x \frac{dx}{2\sqrt{x}} = \sqrt{x} \quad \text{als } x \in (0,1)$$

$$= 0, \text{ als } x < 0$$

$$= 1, \text{ als } x \geq 1.$$

Om uit $F(x)$ te genereren, nemen wij

$$X = F^{-1}(\omega_1) = \omega_1^2.$$

Opgave 10. $X_1 \ldots X_{100} \sim m = 75,\ \sigma^2 = 225$

a) Chebyshev: $P(|\bar{X}_{100} - m| > 3) \leq \frac{Var\,\bar{X}_{100}}{9} = \frac{2.25}{9}$

want $\mathbb{E}\bar{X}_{100} = m$, $Var\,\bar{X}_{100} = \frac{\sigma^2}{100} = 2.25$

b) CLT: $P(|\bar{X}_{100} - m| > 3) = 1 - P(-3 \leq \bar{X}_{100} - m \leq 3) =$

$= 1 - P\left(-\frac{3}{1.5} \leq \underbrace{\frac{\bar{X}_{100} - m}{1.5}}_{Z \sim N(0,1)} \leq \frac{3}{1.5}\right) = 1 - P(-2 \leq Z \leq 2) =$

$= 0.0456$ (tabel)

want standaarde deviatie van $\bar{X}_{100}$ is $\frac{\sigma}{\sqrt{100}} = \frac{\sqrt{225}}{\sqrt{100}} = \frac{15}{10} = 1.5$.

Opgave 11 $x_1, x_2 \ldots x_n$: $f(x) = \begin{cases} \theta^2 x e^{-\theta x}, & x \geq \\ 0, & \text{elders} \end{cases}$

MLS: likelihood:

$$L(\theta) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n \theta^2 x_i e^{-\theta x_i}$$

log-likelihood:

$$\ell(\theta) = \log L(\theta) = \sum_{i=1}^n \log(\theta^2 x_i e^{-\theta x_i}) =$$

$$= 2n \log\theta + \sum_{i=1}^n \log x_i - \theta \sum_{i=1}^n x_i.$$

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{2n}{\theta} - \sum_{i=1}^n x_i = 0 \Rightarrow \theta = \frac{2n}{\sum_{i=1}^n x_i} = \frac{2}{\bar{x}_n}$$

Opgave 12 $X_1, \ldots X_n \sim F(t)$

$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n 1\{X_i \leq t\}$ – empirische verdelingsfunctie

Zij $Y_i = 1\{X_i \leq t\}$. Dan $\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n Y_i$

$Y_i \sim$ Bernoulle (dus 1 met kans p, 0 met kans $1-p$, waar $p = P(X_i \leq t) = F(t)$)

Dan $\mathbb{E}\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbb{E} Y_i = p = P(X_i \leq t) = F(t)$

$Var\,\hat{F}_n(t) = \frac{1}{n^2}\sum_{i=1}^n Var\,Y_i = \frac{1}{n^2}\sum_{i=1}^n p(1-p) = \frac{F(t)(1-F(t))}{n}$

door onafhankelijkheid van X_i's (en dus van Y_i's)

Opgave 13 $X_1 \ldots X_{20} \sim N(\mu, \sigma^2)$.

$H_0: \mu = 35$ tegen $H_1: \mu \neq 35$.

Toetsingsgrootheid: $T_n = \dfrac{\bar{X}_n - 35}{\hat{\sigma}/\sqrt{n}} = \sqrt{n}\,\dfrac{\bar{X}_n - 35}{S_n}$

Uit dataset berekenen:

$\bar{X}_n = 35.07$

$S_n = \sqrt{\dfrac{1}{20-1}\sum_{i=1}^{20}(X_i - \bar{X}_n)^2} = 0{,}166$

Dus $T_n = 1.889$.

Kritieke waarden (uit tabel bij 0.025 (½ van $\alpha = 5\%$ want tweezijdig toetsen) : ± 2.093 (t graden van vrijh. 19).

t-verdeling

Kritiek gebied : $(-\infty, -2.093] \cup [2.093, +\infty)$.

$T_n \notin$ kritiek gebied $\Rightarrow$ Ho niet verwerpen!

Betrouwbaarheidsinterval voor μ: ($\alpha = 5\%$)

$P(-2.093 \leq T_n \leq 2.093) = 95\%$. Dus

$P\left(-2.093 \leq \sqrt{20}\,\dfrac{\bar{X}_n - \mu}{S_n} \leq 2.093\right) = P\left(\dfrac{-2.093 \cdot S_n}{\sqrt{20}} \leq \bar{X}_n - \mu \leq \dfrac{2.093 \cdot S_n}{\sqrt{20}}\right) =$

$= P\left(\bar{X}_n - \dfrac{2.093 S_n}{\sqrt{20}} \leq \mu \leq \bar{X}_n + \dfrac{2.093 S_n}{\sqrt{20}}\right) = P(34.29 \leq \mu \leq 35.85) =$

$= 95\%$

$\Rightarrow$ 95% betr. interval voor μ is: $\boxed{[34.29, 35.85]}$