

Reeksamen i Statistik 2. år

Skriftlig prøve (4 timer)

1. august 2003 kl. 9.00–13.00

Eksamenssættet er på 3 sider.

Alle skriftlige hjælpemidler samt lommeregner er tilladt.

Vægtfordeling: Opgaverne vægtes ens.

Opgave 1

(a) To symmetriske mønter kastes. Lad $X \in \{0, 1, 2\}$ betegne antal mønter som viser krone. Opskriv sandsynlighedsfunktionen for X .

(b) Det i (a) omtalte kast med to mønter udføres 10 gange. Lad S_0 , S_1 og S_2 betegne antal kast blandt de 10 som resulterer i henholdsvis 0, 1 og 2 kroner. Beskriv fordelingen af (S_0, S_1, S_2) , og udregn sandsynligheden for udfaldet

$$\{(S_0, S_1, S_2) = (3, 5, 2)\}.$$

(c) Beskriv den marginale fordeling af S_1 , og udregn middelværdi og standardafvigelse for S_1 .

Opgave 2

(a) Gør rede for at den ved

$$p(x, y) = \frac{1}{2} + 2xy$$

definerede funktion

$$p : [0, 1] \times [0, 1] \rightarrow \mathbf{R}$$

er en sandsynlighedstæthed for en kontinuert fordeling på enhedskvadratet $[0, 1] \times [0, 1]$.

I det følgende betegner (X, Y) en stokastisk variabel med denne fordeling.

(b) Beskriv den marginale fordeling af X .

(c) Gør rede for at X og Y ikke er stokastisk uafhængige.

Opgave 3

Nedenstående tabel viser fordelingen på køn og afdeling for medarbejderne ved en virksomhed i IT-branchen:

	Mænd	Kvinder	I alt
Udvikling	45	7	52
Konsulenter	106	36	142
Salg og markedsføring	16	4	20
Administration og bogholderi	34	22	56

(a) Vi betragter modellen hvor antal kvinder på den enkelte afdeling opfattes som binomialfordelt med det samlede antal medarbejdere på afdelingen som antalsparameter og en sandsynlighedsparameter, der i første omgang afhænger af afdelingen. Estimer parametrene i denne model med angivelse af 95% sikkerhedsgrænser, og kommenter resultatet.

(b) Foretag testet for homogenitet (altså for om de fire sandsynlighedsparametre kan antages at være ens).

(c) Under forudsætning af at testet i spørgsmål (b) fører til forkastelse, undersøg om afvigelsen fra homogenitet alene kan forklares ved den meget skæve kønsproportion i udviklingsafdelingen.

Opgave 4

For 21 støbeprøver af cement er i nedenstående tabel angivet de samhørende værdier af

x = den inverse hærdningstid, og

y = logaritmen til trækstyrken.

Bemærk at forsøget er foretaget således at flere støbeprøver er hærdet det samme antal dage. For eksempel er de fem sidste tal i sidste linie logaritmiske trækstyrker for fem forskellige prøver der er hærdet i 28 dage (idet $0.0357=1/28$).

x	y				
1.0000	1.1139	1.1239	1.0719		
0.5000	1.3404	1.3892	1.3927		
0.3333	1.4742	1.4472	1.3820	1.3838	1.4183
0.1429	1.5105	1.4829	1.5378	1.5198	1.5527
0.0357	1.6212	1.6294	1.6053	1.5527	1.5717

(Kilde A. Hald: Statistiske Metoder, Akademisk Forlag 1968).

(a) De valgte transformationer (invertering af hærdningstiden, log-transformation af trækstyrkerne) er foretaget, fordi man erfaringsmæssigt derved opnår en lineær sammenhæng mellem x og y . Undersøg ved

hjælp af en tegning, om dette ser ud til at være tilfældet. Beskriv forudsætningerne for den tilsvarende simple regressionsmodel med y som afhængig variabel og x som forklarende variabel, og diskuter om forudsætningerne ser ud til at være opfyldt.

(b) Estimer parametrene i den simple regressionsmodel. For hældningens vedkommende ønskes angivelse af 95% konfidensgrænser. Følgende regnestørrelser kan benyttes:

$$S_x = 7.0595$$

$$S_y = 30.1215$$

$$SS_x = 4.41392$$

$$SS_y = 43.73055$$

$$SP_{xy} = 9.11058$$

(c) Den foreliggende gruppering af x -værdierne gør det muligt at teste regressionsmodellen imod en ensidet variansanalysemodel, hvor y -værdierne hørende til samme x -værdi blot antages at have samme middelværdi, uden at denne behøver at afhænge lineært af x . Foretag dette test. Til hjælp for beregningerne oplyses, at residualkvadratsummen i den ensidede variansanalysemodel bliver 0.01681.