

Eksamen i Statistik 2. år

Skriftlig prøve (4 timer)

12. juni 2002 kl. 9.00–13.00

Eksamenssættet er på 3 sider.

Alle skriftlige hjælpemidler samt lommeregner er tilladt.

Vægtfordeling: Opgaverne vægtes ens.

Opgave 1

I denne opgave betegner X_1 og X_2 uafhængige stokastiske variable, som begge er ligefordelt på mængden $\{1, 2, 3, 4\}$. Produktet af disse to variable betegnes $Y = X_1 X_2$.

- (a) Udregn $P(Y = 4)$.
- (b) Udregn $E(Y)$ og $\text{var}(Y)$.
- (c) Opskriv og tegn sandsynlighedsfunktionen for Y .

Opgave 2

Lad R_1 og R_2 være uafhængige stokastiske variable, ligefordelte på enhedsintervallet $[0, 1]$. Vi sætter

$$Y_1 = R_1^3, Y_2 = R_2^3.$$

- (a) Opskriv tætheden for Y_1 .
- (b) Udregn middelværdi og varians for Y_1 .
- (c) Udregn sandsynligheden for at både Y_1 og Y_2 er mindre end 0.5.

Opgave 3

Nedenstående tabel gengiver en del af resultatet af et forsøg til sammenligning af to antidepressive behandlinger. 100 patienter blev ved lodtrækning delt i to lige store grupper, som blev behandlet med hvert sit præparat. For hver behandling er de 50 patienter klassificeret efter graden af uønskede bivirkninger.

Bivirkninger:	ingen	lette	svære
Behandling 1	39	7	4
Behandling 2	20	16	14

Vi betragter modellen hvor rækkesummerne (begge 50) er faste, således at de to rækker opfattes som uafhængige, polynomialfordelte af orden 3 med antalsparameter 50 og hvert sit sæt af sandsynlighedsparametre.

(a) Estimer, med angivelse af 99% sikkerhedsgrænser (OBS: 99, ikke 95), sandsynligheden for at en patient i behandlingsgruppe 1 undgår bivirkninger

(b) Er der forskel på de to behandlinger hvad angår bivirkninger?

(c) Hvis man alene ser på de 41 patienter, for hvilke der er konstateret bivirkninger, er der da forskel på de to behandlinger hvad angår bivirkningernes sværhedsgrad?

Opgave 4

Et maskinværksted modtager præfabrikerede aksler i partier af varierende størrelse. Efter modtagelsen forarbejdes hver enkelt aksel ved at den indsættes i en maskine, hvor den tilpasses ved at det overflødige gods fræses af.

For hvert af 15 fortløbende partier $i = 1, \dots, 15$ har man registret følgende størrelser:

x_i : Antal aksler i partiet.

z_i : Partiets samlede vægt i kg.

y_i : Tidsforbruget i minutter ved forarbejdningen af hele partiet.

i	x_i	z_i	y_i
1	55	5563	1738
2	20	2041	491
3	35	3594	999
4	45	4523	1370
5	40	4082	1150
6	25	2534	684
7	55	5556	1650
8	30	3044	876
9	60	6095	1910
10	45	4561	1380
11	35	3562	995
12	25	2546	660
13	45	4576	1390
14	35	3529	1025
15	30	3056	821

(a) I første omgang ser vi bort fra partiernes vægte (z 'erne). Estimer parametrene (hældning, afskæring og varians) i en simpel regressionsmodel med y som den afhængige variabel og x som forklarende variabel, og

kommenter resultatet. For hældningens vedkommende ønskes angivelse af 95% sikkerhedsgrænser. Til udregningerne kan følgende mellemregningsstørrelser benyttes:

$$55 + 20 + \dots + 30 = 580$$

$$55^2 + 20^2 + \dots + 30^2 = 24450$$

$$1738 + 491 + \dots + 821 = 17139$$

$$1738^2 + 491^2 + \dots + 821^2 = 22051749$$

$$55 \times 1738 + 20 \times 491 + \dots + 30 \times 821 = 733235$$

(b) Kontroller model og beregninger ved hjælp af et x - y plot med den estimerede regressionslinie indtegnet.

(c) I en multipel regressionsmodel med både antal og vægt som forklarende variable, hvor observationen y_i antages at stamme fra en normalfordeling med middelværdi $\alpha + \beta x_i + \gamma z_i$ og varians σ^2 , fås en residualkvadratsum på 10097.0, hvilket er 282.4 mindre end i den simple regressionsmodel fra spørgsmål (a). Udfør test for reduktion af den multiple regressionsmodel til den simple regressionsmodel, og forklar testets konklusion i ord.