KUNGL
TEKNISKA
HÖGSKOLAN

Matematiska institutionen
avd matematisk statistik

TENTAMEN I 5B1503 STATISTIK MED FÖRSÖKSPLANERING
FÖR B OCH K TISDAGEN DEN 2 APRIL 2002 KL 08.00–13.00.

Examinator: Gunnar Englund, tel. 7907416

Tillåtna hjälpmedel: Formel- och tabellsamling i Statistik med försöksplanering. Kalkylator.

Införda beteckningar skall förklaras och definieras. Resonemang och räkningar skall vara så utförliga och motiverade att de är lätta att följa. 20 poäng innebär godkänt resultat.

Resultatet anslås senast tisdagen den 23 april 2002 på matematisk statistiks anslagstavla i entréplanet, Lindstedtsvägen 25. Tentamen kommer att finnas tillgänglig på elevexpeditionen sju veckor efter skrivningstillfället.

Uppgift 1

Antalet fel, X, i en produktenhet, anses vara en poissonfördelad stokastisk variabel med väntevärde 0.2.

a) Vad är sannolikheten att en enhet har fler än 2 fel. (3 p)

b) En köpare av partier bestående av dessa produktenheter använder sig av en internationell standard, ISO 2859, för att kontrollera partiet. Den provtagningsplan som köparen då använder, anger att 80 enheter skall tas ut och partiet godkännas om man finner <u>sammanlagt</u> högst 21 fel. Antag att antalet fel på olika enheter är oberoende av varandra med fördelning enligt ovan. Beräkna sannolikheten att partiet godkänns. Välmotiverade approximationer får användas. (7 p)

Uppgift 2

I många sammanhang har man krav på att mätosäkerheter skall anges tillsammans med en skattning av en storhet. Detta görs ofta genom att ange ett konfidensintervall för det okända värdet på storheten ifråga.

Vid bestämning av kolhalt i prover utförde ett laboratorium 6 mätningar av halten med följande resultat.

1.23 1.28 1.11 1.19 1.32 1.25 (%)

Observationerna anses komma från en normalfördelning med väntevärde m och varians σ^2, där m är den verkliga kolhalten.

a) Ange en skattning av den verkliga kolhalten och osäkerheten i skattningen genom att beräkna ett 95 % konfidensintervall för m. σ antas okänd i denna uppgift. (3 p)

b) σ är ett mått på mätmetodens precision. Det uppges att σ skall vara 0.05. Testa denna hypotes på 5 % signifikansnivå genom att t.ex. beräkna ett 95 % konfidensintervall för σ. Skall hypotesen σ=0.05 förkastas eller inte? (3 p)

c) Antag att σ=0.05. Hur många observationer skall man göra om man vill att ett 95 % konfidensintervall för m skall ha längd högst 0.05? (4 p)

Uppgift 3

Tre termometrar används regelbundet i ett laboratorium. För att kontrollera den relativa noggrannheten hos termometrarna placerades dessa i slumpmässig ordning i en cell som höll 0 ^{o}C. Varje termometer placerades i cellen fyra gånger, med följande resultat.(^{o}C).

Termometer	1	2	3
	0.10	-0.20	0.90
	0.90	0.80	0.20
	-0.80	-0.30	0.30
	-0.20	0.60	-0.30

Hjälpsumma:$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = 3.3267$

a) Ange lämplig statistisk modell samt testa på 5%-nivån om termometrarna mäter likvärdigt. (6 p)

b) Reducera eventuellt modellen samt beräkna ett 95% konfidensintervall för den första termometerns systematiska fel. (4 p)

Uppgift 4

Enligt Hookes lag är förlängningen y av en fjäder en linjär funktion av belastningen x. Vid konstruktionen av en våg har man använt sig av denna princip. För att kalibrera vågen mätte man förlängningen y av fjädern för var och en av 9 olika precisionsbestämda vikter x_i, $i = 1, 2, \ldots, 9$. Följande värden erhölls:

x_i	4	5	6	7	8	9	10	11	12
y_i	6.2	7.1	7.9	8.9	10.8	12.3	12.7	13.6	15.0

Man beräknade följande storheter.

$$\sum y_i = 94.5 \quad \sum y_i^2 = 1069.65 \quad \sum y_i\big(x_i - \bar{x}\big) = 67.7 \quad \sum \big(x_i - \bar{x}\big)^2 = 60$$

a) Ansätt en enkel regressionsmodell $y_i = \alpha + \beta x_i + \varepsilon_i$, där $\varepsilon \in N(0, \sigma)$ och skatta α, β och σ^2. (3 p)

b) Bestäm ett 95% konfidensintervall för β. (3 p)

c) Antag att man för ett okänt värde på x, säg x_0, mätt motsvarande y-värde till 11.4. Beräkna en skattning av x_0 och ett approximativt 95% konfidensintervall för x_0. (4 p)

Uppgift 5

I ett försök varierades faktorerna A, B, C och D på 2 nivåer vardera (+ respektive –). Följande observationer erhölls.

A	B	C	D	y
$-$	$-$	$-$		268
$+$	$-$	$-$		252
$-$	$+$	$-$		414
$+$	$+$	$-$		342
$-$	$-$	$+$		218
$+$	$-$	$+$		158
$-$	$+$	$+$		284
$+$	$+$	$+$		304

Man valde försökspunkterna så att kopplingen $I - ABCD$ erhölls. Om försöket betraktas som ett 2^3-försök i faktorerna A, B och C fås följande skattningar:
$\widehat{A} = -16, \widehat{B} = 56, \widehat{AB} = 3, \widehat{C} = -39, \widehat{AC} = 6$ och $\widehat{BC} = -3$.

a) Ange samtliga kopplingar i försöket. (1 p)

b) Ange de utelämnade försökspunkterna för faktorn D i tabellen ovan. (1 p)

c) Av erfarenhet tror man sig veta att samspelen kan försummas. Utnyttja detta för att avgöra vilka huvudfaktorer som kan anses vara signifikanta. 5 % signifikansnivå. (5 p)

d) Anta som i c) att faktorerna är additiva. Ge ett 95 % konfidensintervall för förväntat värde då alla faktorerna är på sin +nivå. (3 p)

KUNGL
TEKNISKA
HÖGSKOLAN

Matematiska institutionen
avd matematisk statistik

LÖSNING FÖR TENTAMEN I 5B1503 STATISTIK MED FÖRSÖKSPLANERING FÖR B OCH K TISDAGEN DEN 2 APRIL 2002 KL 08.00–13.00.

Uppgift 1

a) Sätt X=antalet fel. $X \in \text{Po}(0.2)$. Alltså är $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.99885 = 0.00115$. Detta fås ut tabell.

b) Sätt X_i = antal fel på i:te komponenten, $i = 1, 2, \ldots, 80$. $X_1, X_2, \ldots$ är oberoende alla Po(0.2). Summan är då Po($80 \cdot 0.2$)=Po(16), som sin tur är approximativt N(16,16). Detta ger att

$$P(Y \leq 21) \approx \Phi(\frac{21 - 16}{\sqrt{16}}) = \Phi(1.25) = 0.894$$

Alternativt kan man använda cgs; man får samma resultat.
Anm. Exakt värde är 0.911.

Uppgift 2

a) Beräkningar ger $\bar{x} = 1.23$ och stickprovsvarians $s^2 = 0.0054$. Ett 95 % konfidensintervall för m ges av $\bar{x} \pm t_{0.025}(6-1)s/\sqrt{6} = \underline{1.230 \pm 0.077}$.

b) Ett konfidensintervall för σ^2 ges av

$$(\frac{(n-1)s^2}{\chi^2_{0.025}(n-1)}, \frac{(n-1)s^2}{\chi^2_{0.975}(n-1)})$$

$\chi_{0.025}(5) = 12.8$ och $\chi_{0.975}(5) = 0.83$, vilket ger intervallet $\big(0.0021, 0.0325\big)\cdot$, varför ett 95 % konfidensintervall för σ blir $\underline{\big(0.045, 0.18\big)}$. Hypetosen σ=0.05 kan inte förkastas eftersom 0.05 ligger i konfidensintervallet.

c) Villkoret ger att $2 \cdot \lambda_{0.025}\sigma/\sqrt{n} \leq 0.05$ dvs $2 \cdot 1.96 \cdot 0.05/\sqrt{n} \leq 0.05$ vilket ger $\sqrt{n} \geq 3.92$ vilket ger $\underline{n \geq 16}$ (observera n heltal).

Uppgift 3

a) Ensidig variansanalys. Låt y_{ij}=observation nr j från termometer nr i.
Då fås $\bar{y}_{1.} = 0$ $\bar{y}_{2.} = 0.225$ $\bar{y}_{3.} = 0.275$ och $\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 = 0.0429$. Härav fås variansanalystabellen

Källa	Fg	Kvs	Mkvs
Mellan termometrar	2	0.1717	0.0858
Inom termometrar	9	3.1550	0.3506
Totalt	11	3.3267	

Testkvot $F_{termometer} = \frac{0.0858}{0.3506} = 0.2448 < F_{0.05}(2,9) = 2.26$ och således kan man inte förkasta hypotesen att termomterarna mäter likvärdigt.

b) Vi reducerar modellen och antar att termometrarna är likvärdiga. Ny σ^2-skattning blir $s^2 = 3.3267/11$ och konfidensintervall för systematiska felet $\bar{y}_{..} \pm t_{0.025}(11)s/\sqrt{12} = \underline{0.17 \pm 1.16}$. Man kan också tänka sig intervallet $\bar{y}_{1.} \pm t_{0.025}(11)s/\sqrt{4}$.

Uppgift 4

a) Med sedvanliga beteckningar har vi

$$\widehat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{67.7}{60} = 1.128 \quad \widehat{\alpha} = \bar{y} - \widehat{\beta}\bar{x} = 10.5 - 1.1283 \cdot 8 = 1.473 \text{ och} \tag{1}$$

$$\widehat{\sigma}^2 = \frac{1}{9-2}\left(S_{yy} - S_{xy}^2/S_{xx}\right) = \frac{1}{7}(77.4 - 67.7^2/60) = 0.1445 \tag{2}$$

eftersom $S_{yy} = \sum y_i^2 - n \cdot \bar{y}^2 = 1069.65 - 7 \cdot 10.5^2 = 77.4$

b) Konfidensintervallet ges av $\widehat{\beta} \pm t_{0.025}(9-2)\frac{\widehat{\sigma}}{\sqrt{S_{xx}}} = 1.128 \pm 2.36\sqrt{0.1445/60} = 1.128 \pm 0.116$

c) Vi löser ut x_0 ur regressionslinjen; $x_0 = \frac{y_0 - \widehat{\alpha}}{\widehat{\beta}} = \frac{11.4 - 1.473}{1.128} = 8.800$. Konfidensintervall ges av (se FS)

$$\widehat{x}_0 \pm t_{0.025}(7)\frac{\widehat{\sigma}}{\widehat{\beta}}\sqrt{1 + \frac{1}{n} + \frac{(\widehat{x}_0 - \bar{x})^2}{S_{xx}}} = 8.800 \pm 0.844$$

Uppgift 5

a) Kopplingarna blir $A-BCD, B-ACD, C-ABD, D-ABC, AB-CD, AC-AD$ och $AD-BC$ förutom den angivna $I - ABCD$.

b) Eftersom produkten av tecknen för samtliga kolumner skall vara – fås försökspunkterna $+ - - + - + + -$.

c) I verkligheten är de beräknade skattningarna skattningar av $A + BCD, B + ACD$ osv. Skattningar av $D + ABC$ och $I + ABCD$ är ej beräknade. De blir på sedvanligt sätt (om samspelen försummas)
$\widehat{D} = \frac{268-252-414+342-218+158+284-304}{8} = 17$ och $\widehat{I}$ = medelvärden av data = 280. Om samspelen försummas kan skattningar av $AB + CD, AC + BD$ och $AD + BC$ användas för att skatta

$\widehat{\sigma}^2$. Vi får $\widehat{\sigma}^2 = \dfrac{8 \cdot (3^2 + 6^2 + 3^2)}{3} = 144$. Konfidensintervall för effekter ges av skattning $\pm \widehat{t}_{0.025}(3)\sigma/\sqrt{8}$= skattning ± 13.5. Alla huvudfaktorer signifikanta.
d) Konfidensintervall för $I + A + B + C + D$ söks. Punktskattning blir 280-16+56-39+17=298. Medelfel är $\widehat{\sigma}\sqrt{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}} = \widehat{\sigma}\sqrt{5/8} = 9.49$.
Konfidensintervallet: skattning $\pm t_{0.025}(3)\cdot$medelfel dvs 298 ± 30.19.