Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1923 SANNOLIKHETSTEORI OCH STATISTIK, SAMT TENTAMEN I SF1935 SANNOLIKHETSTEORI OCH STATISTIK MED TILLÄMPNING INOM MASKININLÄRNING MÅNDAG 25 MAJ 2026 KL 8.00–13.00.

*Kursledare*: Björn-Olof Skytt, 08-790 86 49

*Tillåtna hjälpmedel*: Formel- och tabellsamling i Matematisk statistik (utdelas vid tentamen), miniräknare.

Tentamen består av två delar, benämnda del I och del II.
Del I består av uppgifterna 1–12 och varje korrekt svar ger 1 poäng. På denna del ska endast svar anges i form av val av ett av de möjliga svarsalternativen på de uppgifter där svarsalternativ finns, annars skall svaret ges som ett numeriskt värde med minst tre värdesiffrors noggrannhet!. Svaren ska anges på svarsblanketten (utdelas vid tentamen). Studenter som är godkända på kontrollskrivningen behöver ej besvara uppgift 1–3, utan får tillgodoräkna sig dessa tre uppgifter (i svarsblanketten anges då ordet ”Bonus”). Studenter som är godkända på den andra datorlaborationen behöver ej besvara uppgift 12, utan får tillgodoräkna sig denna uppgift (i svarsblanketten anges då ordet ”Bonus”). Dessa tillgodoräknanden gäller för den här tentamen och vid omtentamen i augusti 2026. Gränsen för godkänt är 9 poäng. Möjlighet att komplettera ges för tentander med 8 poäng.

Del II består av uppgifterna 13–16 och varje korrekt lösning ger 10 poäng. Del II rättas bara för studenter som är godkända på eller får komplettera del I och poäng på del II krävs för högre betyg än E. På denna del ska resonemang och uträkningar vara så utförliga och väl motiverade att de är lätta att följa. Införda beteckningar ska förklaras och definieras samt numeriska svar ska anges med *minst tre* värdesiffrors noggrannhet. Studenter som är godkända på den andra datorlaborationen får dessutom tre bonuspoäng på del II. Dessa bonuspoäng gäller för den här tentamen och vid omtentamen i augusti 2026.

Tentamen kommer att vara rättad inom tre arbetsveckor från skrivningstillfället och kommer att finnas tillgänglig på studentexpeditionen minst sju veckor efter skrivningstillfället.

# Del I

### Uppgift 1

Händelserna $A$ och $B$ är disjunkta och händelserna $A$ och $C$ är oberoende.

$P(A) = 0.4$, $P(B) = 0.3$, $P(C) = 0.5$ och $P(B|C^*) = 0.2$.

($C^*$ är komplementet till $C$.)

Beräkna $P(A^* \cap B^* \cap C)$.

## Uppgift 2

Låt fördelningsfunktionen för den stokastiska variabln $X$ vara

$$F_X(x) = \begin{cases} 0 & \text{om } x \leq 0. \\ c \cdot x^2 & \text{där c är en konstant och } 0 \leq x \leq 2 \\ 1 & \text{om } x \geq 2 \end{cases}$$

Beräkna $E(X^2)$.
OBS! Svaret skall bli ett reellt tal, dvs ej som en funktion av $c$.

## Uppgift 3

Vid ett strömavbrott i fritidshuset går sommargästen till sin låda med batterier. I lådan ligger 5 batterier varav 3 inte fungerar. Sommargästen tar i mörkret på måfå ut två batterier. Vad är sannolikheten att minst ett av dem fungerar?

A: ca 0.48

B: ca 0.60

C: ca 0.64

D: ca 0.70

## Uppgift 4

I en produktionsprocess blir enheter felaktiga oberoende av varandra och alla med sannolikhet 0.007. Man tillverkar 1000 enheter. Beräkna sannolikheten att högst 5 av dessa är felaktiga. Använd tillåten approximation.

A: ca 0.22

B: ca 0.30

C: ca 0.35

D: ca 0.45

## Uppgift 5

$X \in Po(4)$ och $Y \in Po(2)$. $X$ och $Y$ är oberoende. $Z = min(X, Y)$.

Bestäm $P(Z \leq 3)$.

A: ca 0.08

B: ca 0.37

C: ca 0.63

D: ca 0.92

## Uppgift 6

$X, Y$ och $Z$ är oberoende normalfördelade stokastiska variabler, $E(X) = 2$, $E(Y) = 1$ och $E(Z) = 0$. Alla har varians 2. Beräkna $P(4X - 3Y > 5Z)$.

A: ca 0.52

B: ca 0.69

C: ca 0.80

D: ca 0.94

## Uppgift 7

Vi har de två oberoende väntevärdesriktiga skattningarna $\hat{\theta}_{1_{obs}}$ och $\hat{\theta}_{2_{obs}}$ som har standardavvikelserna $D(\hat{\theta}_1) = 0.4$ och $D(\hat{\theta}_2) = 0.3$.

Vi bildar $\theta^*_{1_{obs}} = 0.3\hat{\theta}_{1_{obs}} + 0.7\hat{\theta}_{2_{obs}}$ och $\theta^*_{2_{obs}} = 0.4\hat{\theta}_{1_{obs}} + 0.6\hat{\theta}_{2_{obs}}$.

Vilket av nedanstående påståenden är sant?

A: $\theta^*_{1_{obs}}$ är den effektivaste skattningen av $\theta$.

B: $\theta^*_{2_{obs}}$ är den effektivaste skattningen av $\theta$.

C: Bägge skattningarna är lika effektiva.

D: Man kan inte avgöra vilken av skattningarna som är effektivast, eftersom minst en av dem inte är väntevärdesriktig.

### Uppgift 8

I en veckotidning beskrivs en bantningskur som påstås ge en viktminskning av ungefär 7 kg på 14 dagar. En grupp på åtta personer bestämmer sig för att pröva kuren. Deras vikter (i kg) före och efter genomgången kur ges nedan:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Före | 108 | 88 | 82 | 103 | 98 | 100 | 90 | 85 |
| Efter | 110 | 85 | 78 | 101 | 91 | 99 | 90 | 82 |

Vikterna kan anses vara observationer av oberoende normalfördelade stokastiska variabler. Det är dock inte rimligt att anta att samtliga personer har samma förväntade vikt före kuren, eller att de har samma förväntade vikt efter kuren.

Ange undre gränsen för ett tvåsidigt konfidensintervall av grad 95% för den förväntade viktminskningen efter genomgången kur.

A: ca -8.5

B: ca -6.6

C: ca 0.0

D: ca 0.4

### Uppgift 9

Antag att $X \in Bin(n, p)$. Vi gör $n$= 500 försök och får $x = 40$.
Ange övre gränsen för det ensidigt uppåt begränsade konfidensintervallet $I_p$ för $p$ då den approximativa konfidensgraden är 90%.

### Uppgift 10

Antag att $T \in Exp(\lambda)$. Vi vill testa $H_0 : \lambda = 0.01$ mot alternativet $H_1 : \lambda = 0.1$. Vi förkastar $H_0$ till förmån för $H_1$ om vi får ett utfall $t \leq 30$. Vi får en observation: $t = 15$.
Bestäm P-värdet.

A: ca 0.14

B: ca 0.24

C: ca 0.78

D: ca 0.95

## Uppgift 11

Vid en Folk- och Bostadsräkningen, som gjordes för c:a 25 år sedan, hade en miljonpopulation av individer klassificerats i fyra klasser A, B, C, D, vars relativa storlekar var 20%, 30%, 10%, 40% respektive. 10 år senare klassificerades ett slumpmässigt urval om 100 individer på samma sätt. Av dessa tillhörde 13, 37, 17, 33 i nämnd ordning ovanstående klasser.
Utför ett test för att undersöka om proportionerna av klasserna i populationen hade förändrats under de 10 åren. Låt nollhypotesen $H_0$ vara ”relativa storlekarna är oförändrade”.
Vilket påstående stämmer?

A: $H_0$ kan förkastas på risknivån 1%, men inte på risknivån 0.1%.

B: $H_0$ kan förkastas på både risknivån 1% och risknivån 0.1%.

C: $H_0$ kan varken förkastas på risknivån 1% eller risknivån 0.1%.

D: $H_0$ kan förkastas på risknivån 0.1%, men inte på risknivån 1%.

## Uppgift 12

Beräkna minsta-kvadrat skattningen $\theta^*_{obs}$ av $\theta$ då man erhållit observationerna $0.15, 0.40, 0.28$, och $0.70$ på en stokastisk variabel med täthetsfunktionen

$$f_X(x) = \begin{cases} \theta x^{\theta-1} & \text{om } 0 \le x \le 1 \\ 0 & \text{annars} \end{cases}$$

A: ca 0.28

B: ca 0.62

C: ca 1.61

D: ca 3.61

**Var god vänd!**

# Del II

### Uppgift 13

Lisa har fem kort, av vilka ett är rött på båda sidorna, ett är vitt på båda sidorna och de tre andra är vita på den ena sidan och röda på den andra sidan. Pelle, en kompis till Lisa, drar ett av korten på måfå, de tittar tillsammans på ena sidan och ser att den är röd. Pelle satsar 300 kr på att den andra sidan också är röd och vill att Lisa skall satsa på att den andra sidan är vit. Lisa antar utmaningen och satsar $L$ kr. De tittar tillsammans på den andra sidan och den som vinner får hela banken, dvs $300 + L$ kr (vinsten här är då den andres insats).

(a) Vad är sannolikheten att Pelle vinner respektive att Lisa vinner? (6 p)

(b) Bestäm $L$ sådan att spelet ska bli rättvist, dvs sådan att den förväntade vinsten är lika för dom båda? (4 p)

### Uppgift 14

Två personer (oberoende av varandra) kastar varsin rättvisa sexsidiga tärning tills de båda får en 6:a.

(a) Vad är sannolikheten att dessa två personer måste kasta tärningen lika många gånger? (4 p)

*Ledning*: $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, om $0 < x < 1$.

(b) Vad är sannolikheten att den ena personen måste kasta sin tärning dubbelt så många gånger som den andra personen? (4 p)

(c) (*Fortsättning på (a)*) Antag att vi har $m$ personer, $m \geq 2$. Sannolikheten att det måste kastas lika många gånger för samtliga dessa personer tills dom alla får en 6:a kan skrivas på formen $1/(k^m - (k-1)^m)$, där $k$ är ett positivt heltal. Bestäm $k$ och redogör för att det stämmer för $m = 3$. (2 p)

### Uppgift 15

Antal telefonsamtal som kan göras samtidigt till en telefonist vid en växelstation kan uppfattas som en stokastisk variabel $X$ med täthetsfunktionen

$$f_X(x) = \begin{cases} 0 & \text{om } x < 0. \\ \dfrac{x^3 e^{-x/\beta}}{6\beta^4} & \text{om } x \geq 0. \end{cases}$$

Man kan visa att $E(X) = 4\beta$ och $E(X^2) = 20\beta^2$. En fördelning av denna typ kallas ibland för en *Erlangfördelning*. Antag att vi har $n$ observationer $x_1, \ldots, x_n$ på antal telefonsamtal till oberoende telefonister. Dvs $x_1, \ldots, x_n$ uppfattas som utfall av oberoende stokastiska variabler med ovanstående fördelning, där $\beta$ är en positiv men i övrigt okänd parameter. Härled ML-skattningen $\beta^*_{obs}$ av $\beta$ och bestäm sedan medelfelet för denna skattning om vi fått utfallen $x_1 = 4$, $x_2 = 5$, $x_3 = 1$ och $x_4 = 10$ (dvs utgående från dessa fyra obersvationer). (10 p)

**Var god vänd!**

**Uppgift 16**

Smältpunkten för två legeringar $A$ och $B$ mäts med hjälp av en termometer. Antag att motsvarande mätningar är oberoende och normalfördelade $N(\mu_A, \sigma_A)$ respektive $N(\mu_B, \sigma_B)$, där $\sigma_A = 2.5$ respektive $\sigma_B = 3.5$ är kända. Vi prövar nollhypotesen att smältpunkterna sammanfaller ($H_0 : \mu_A = \mu_B$) mot alternativhypotesen att de skiljer sig ifrån varandra ($H_1 : \mu_A \neq \mu_B$), $\alpha = 0.05$. Konfidensintervallsmetoden används här för att testa hypotesen $H_0$ och det 95%-iga konfidensintervallet för $\mu_A - \mu_B$ beräknas till:

$$I_{\mu_A - \mu_B}(0.95) = \big(\bar{x}_A - \bar{x}_B \pm \text{ felmarginal}\big) = \big(-5.076, 0.256\big),$$

där $\bar{x}_A = 179.21$, $\bar{x}_B = 181.62$. Bestäm testets styrka då $\mu_A = \mu_B + 5$. (10 p)

**Lycka till!**

Avd. Matematisk statistik

KTH Matematik

# LÖSNINGSFÖRSLAG TENTAMEN I SF1923/SF1935 SANNOLIKHETSTEORI OCH STATISTIK, MÅNDAG 25 MAJ 2026 KL 8.00–13.00.

## Del I, Svar.

1. 0.1

2. 2

3. D

4. B

5. D

6. B

7. B

8. C

9. 0.0955

10. A

11. C

12. B

**Del I, Lösningsförslag.**

### Uppgift 1

$P(C) = P(C \cap A) + P(C \cap B) + P(C \cap A^* \cap B^*)$ eftersom $A$ och $B$ är disjunkta.

$P(C \cap A) = P(A)P(C) = 0.4 \cdot 0.5 = 0.2$ eftesom $A$ och $C$ är oberoende.

$$0.2 = P(B|C^*) = \frac{P(B \cap C^*)}{P(C^*)} = \frac{P(B \cap C^*)}{1 - P(C)} \Rightarrow P(B \cap C^*) = 0.5 \cdot 0.2 = 0.1$$

Eftersom $P(B \cap C^*) + P(B \cap C) = P(B) = 0.3$ så gäller att $P(B \cap C) = 0.2$

Nu har vi $0.5 = 0.2 + 0.2 + P(C \cap A^* \cap B^*)$ vilket ger att $P(C \cap A^* \cap B^*) = 0.1$

### Uppgift 2

$$\int_0^2 f_X(x)dx = 1 \Rightarrow \int_0^2 c \cdot 2xdx = [cx^2]_0^2 = 4c = 1 \Rightarrow c = \frac{1}{4}$$

$$E(X^2) = \int_0^2 x^2 \cdot f_X(x)dx = \int_0^2 x^2 \cdot \frac{2x}{4}dx = \left[\frac{2x^4}{16}\right]_0^2 = 2$$

### Uppgift 3

Här har vi hypergeometrisk fördelning.

$$\frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} + \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} = \frac{3 \cdot 2 + 1 \cdot 1}{\frac{5 \cdot 4}{2 \cdot 1}} = 0.7$$

### Uppgift 4

Låt $X$ vara antalet felaktiga enheter. Då är $X$ antalet gånger händelsen att felaktig enhet tillverkats har uppkommit i 100 oberoende försök och således är $X \in$Bin(1000,0.007). Vi kan inte använda tabell 6 och därför approximerar vi. Sannolikheten $p < 0.1$ och Poissonapproximation är tillåten, $X$ är approximativt Po($100 \cdot 0.07$)=Po(7). Vi erhåller $P(X \leq 5) = 0.30071$ i tabell 5.

### Uppgift 5

$P(Z \leq 3) = P(min(X, Y) \leq 3) = 1 - P(min(X, Y) > 3) = 1 - P(X > 3)P(Y > 3) =$
$1 - (1 - P(X \leq 3))(1 - P(Y \leq 3)) =$ [se tab 5] $= 1 - (1 - 0.43347)(1 - 0.85712) = 0.92$

### Uppgift 6

Vi har att $P(4X - 3Y > 5Z) = P(5Z - 4X + 3Y \leq 0)$. Sätt $U = 5Z - 4X + 3Y$. Vi får att $E(5Z - 4X + 3Y) = 5 \cdot 0 - 4 \cdot 2 + 3 \cdot 1 = -5$ och $V(U) = V(5Z - 4X + 3Y) = 25V(Z) + 16V(X) + 9V(Y) = 100$, dvs $D(Z) = 10$ Alltså får vi att $U \in$N(-5,10) och
$P(U \leq 0) = \Phi\left(\frac{0-(-5)}{10}\right) = \Phi(0.5) = 0.6915$

**Uppgift 7**

Vi har de två oberoende väntevärdesriktiga skattningarna $\hat{\theta}_{1_{obs}}$ och $\hat{\theta}_{2_{obs}}$ som har standardavvikelserna $D(\hat{\theta}_1) = 0.4$ och $D(\hat{\theta}_2) = 0.3$

Vi bildar $\theta^*_{1_{obs}} = 0.3\hat{\theta}_{1_{obs}} + 0.7\hat{\theta}_{2_{obs}}$ och $\theta^*_{2_{obs}} = 0.4\hat{\theta}_{1_{obs}} + 0.6\hat{\theta}_{2_{obs}}$

$V(\theta^*_1) = V(0.3\hat{\theta}_1 + 0.7\hat{\theta}_2) = 0.3^2 \cdot 0.4^2 + 0.7^2 \cdot 0.3^2 = 0.0585$

$V(\theta^*_2) = V(0.4\hat{\theta}_1 + 0.6\hat{\theta}_2) = 0.4^2 \cdot 0.4^2 + 0.6^2 \cdot 0.3^2 = 0.0580$

Alltså är $\theta^*_{2_{obs}}$ den effektivaste skattningen av de två.

**Uppgift 8**

a) Beteckna person nr. $i$:s vikt före resp. efter genomgången kur med $x_i$ resp. $y_i$, för $i = 1, \ldots, 8$. Vikterna är observationer av stokastiska variabler $X_1, \ldots, X_8$ resp. $Y_1, \ldots, Y_8$, och lämpliga modellantaganden är att $X_i \sim N(m_i, \sigma_1)$ och $Y_i \sim N(m_i - \Delta, \sigma_2)$, för $i = 1, \ldots, 8$, där $\Delta$ är den förväntade viktminskningen, och alla parametrar är okända (standardmodellen “stickprov i par”). Bilda de parvisa differenserna $Z_i = X_i - Y_i$, och motsvarande observerade värden $z_i = x_i - y_i$, för $i = 1, \ldots, 8$. Då är $Z_i \sim N(\Delta, \sigma_z)$, där $\sigma_z = \sqrt{\sigma_1^2 + \sigma_2^2}$. De observerade värdena $z_1, \ldots, z_8$ är

$$-2 \quad 3 \quad 4 \quad 2 \quad 7 \quad 1 \quad 0 \quad 3$$

Eftersom $\sigma_z$ är okänt så är

$$I_\Delta = \bar{z} \pm t_{\alpha/2}(n-1)\frac{s_z}{\sqrt{n}}$$

ett konfidensintervall för $\Delta$ av grad $1 - \alpha$. I vårt fall gäller

$$\bar{z} = 2.25 \qquad s_z = 2.7124 \qquad n = 8 \qquad t_{0.025}(7) = 2.36$$

vilket ger

$$I_\Delta = 2.25 \pm 2.36 \cdot \frac{2.7124}{\sqrt{8}} = \underline{2.25 \pm 2.26 = (-0.0132, 4.51)}.$$

**Uppgift 9**

$X \in Bin(n, p) = Bin(500, p)$
Vi får då konfidensintervallets övre gräns till

$$p^*_{obs} + \sqrt{\frac{p^*_{obs} \cdot (1 - p^*_{obs})}{n}} \cdot \lambda_\alpha = \frac{40}{500} + \sqrt{\left(\frac{\frac{40}{500} \cdot (1 - \frac{40}{500})}{500}\right)} \cdot \lambda_{0.10} = \frac{40}{500} + \sqrt{\left(\frac{\frac{40}{500} \cdot (1 - \frac{40}{500})}{500}\right)} \cdot 1.2816 = 0.0955$$

**Uppgift 10**

P-värdet = sannolikheten att förkasta $H_0$ om $H_0$ är sann givet att vi fått det utfall vi fått eller ett utfall som avviker ännu mer mo $H_1$ från det det “borde”vara enligt $H_0$.

Vi ska alltså ta fram $P(T < 15)$ om $\lambda = 0.01$. $P(T < 15) = \int_0^{15} f_T(t)dt = \int_0^{15} 0.01e^{-0.01t}dt = [-e^{0.01t}]_0^{15} = 1 - e^{-0.15} = 0.14$

## Uppgift 11

Här har vi test av given fördelning.

$$Q = (13-20)^2/20 + (37-30)^2/30 + (17-10)^2/10 + (33-40)^2/40 = 490/48 = 10.2.$$

Om hypotesen $H_0$ "relativa storlekarna är oförändrade" är sann så är 10.2 observation av (approximativt) $\chi^2(3)$. Hypotesen förkastas för stora värden på $Q$. Ur tabell erhålls $10.2 < \chi^2_{0.01}(3) = 11.3 < \chi^2_{0.001}(3) = 16.3$. varför $\underline{H_0 \text{ varken kan förkastas på risknivån } 0.1\%}$ eller 1 %.

## Uppgift 12

$$E(X) = \int_0^1 x \cdot \theta \cdot x^{\theta-1} dx = \theta \left[ \frac{x^{\theta+1}}{\theta+1} \right]_0^1 = \frac{\theta}{\theta+1}$$

Kvadratsumman

$$Q = \sum_{i=1}^{4} (x_i - \frac{\theta}{\theta+1})^2$$

skall minimeras.

$$Q' = -2 \sum_{i=1}^{4} (x_i - \frac{\theta}{\theta+1}) \cdot \frac{(\theta+1)-\theta}{(\theta+1)^2}$$

$Q' = 0$ när $\sum_{i=1}^4 x_i = 4\frac{\theta}{\theta+1}$

MK-skattningen av $\theta$ uppfyller då $\frac{\theta^*_{obs}}{\theta^*_{obs}+1} = \bar{x}$, vilket innebär att $\theta^*_{obs} = \frac{\bar{x}}{1-\bar{x}}$. Med siffror insatta erhålls $\bar{x} = 0.3825$ och alltså $\theta^*_{obs} = 0.6194$.

**Del II, Lösningsförslag.**

## Uppgift 13

(a) Låt händelserna $A$, $B$ och $C$ vara:
$A = \{$ett kort med rött på båda sidor$\}$,
$B = \{$att få upp en röd sida$\}$,
$C = \{$ett kort med vitt på ena och rött på den andra sidan$\}$.

Då har vi
$P(\text{att Pelle vinner}) = P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{(1)\cdot(1/5)}{(5/10)} = \frac{2}{5}$
och
$P(\text{att Lisa vinner}) = P(C|B) = \frac{P(B|C)P(C)}{P(B)} = \frac{(1/2)\cdot(3/5)}{(5/10)} = \frac{3}{5}$.

(b) Från (a), samt om Pelle satsar 300 kr och Lisa satsar $L$ kr, så är den förväntade vinsten för Pelle och Lisa lika med $L\,(2/5) - 300\,(3/5)$ respektive $300\,(3/5) - L\,(2/5)$. Därmed, har vi följande ekvation:

$$L\,(2/5) - 300\,(3/5) = 300\,(3/5) - L\,(2/5)\,,$$

ty den förväntade vinsten skall vara lika för dom båda, för att spelet ska bli rättvist. Lösningen till den ekvationen blir $L = 450$ (vilket även innebär att den förväntade vinsten för dom båda kommer att vara lika med noll då). Dvs Lisa ska satsa 450 kr.

## Uppgift 14

Låt $X_1$ och $X_2$ vara två oberoende stokastiska variabler som anger antal kast (inklusive) tills den första 6:an dyker upp för det första respektive för den andra personen. Då vet vi:

$$X_1 \in ffg(p) \quad \text{och} \quad X_2 \in ffg(p)\,,$$

där $p = 1/6$.

(a) Vi är ute efter $P(X_1 = X_2)$. Vi har:

$$P(X_1 = X_2) = \sum_{k=1}^{\infty} P(\{X_1 = k\} \cap \{X_2 = k\}) = |\mathit{oberoendet}| = \sum_{k=1}^{\infty} P(X_1 = k)P(X_2 = k) =$$

$$= \sum_{k=1}^{\infty} \left((1-p)^{k-1}p\right)^2 = p^2 \sum_{k=1}^{\infty} ((1-p)^2)^{k-1} = |\text{Ledning, ty } 0 < (1-1/6)^2 < 1| =$$

$$= p^2 \frac{1}{1-(1-p)^2} = \frac{p^2}{2p-p^2} = \frac{p}{2-p} = |p = 1/6| = \frac{1/6}{2-1/6} = \frac{1}{11} \simeq 0.0909$$

(b) Här är vi ute efter $P(X_1 = 2X_2) + P(2X_1 = X_2)$. Vi har:

$$P(X_1 = 2X_2) = \sum_{k=1}^{\infty} P(\{X_2 = k\} \cap \{X_1 = 2k\}) = |oberoendet| =$$

$$= \sum_{k=1}^{\infty} P(X_2 = k)P(X_1 = 2k) = \sum_{k=1}^{\infty}(1-p)^{k-1}p(1-p)^{2k-1}p = p^2 \sum_{k=1}^{\infty}(1-p)^{3k-3+1} =$$

$$= p^2(1-p) \sum_{k=1}^{\infty} \left((1-p)^3\right)^{k-1} = |\text{Ledning, ty } 0 < (1-1/6)^3 < 1| =$$

$$= \frac{p^2(1-p)}{1-(1-p)^3} = \frac{p(1-p)}{3-3p+p^2} = |p = 1/6| = \frac{(1/6)\cdot(5/6)}{3-3/6+1/36} = \frac{5}{91} \simeq 0.0549$$

Alltså blir svaret

$$2 \cdot \frac{5}{91} = \frac{10}{91} = 0.10989$$

(c) Inte så svårt att inse (från (a)) att $k = 6$, dvs inse att (för $m = 2$)

$$\frac{1}{k^2 - (k-1)^2} = \frac{1}{11},$$

för $k = 6$, tex man kan lösa ekvationen $k^2 - (k-1)^2 = 11$ för det.

Vu har nu för $m = 3$:

$$\frac{1}{6^3 - 5^3} = \frac{1}{91}.$$

Låt oss visa att detta stämmer. Låt $X_1$, $X_2$ och $X_3$ vara tre oberoende stokastiska variabler som anger antal kast (inklusive) tills den första 6:an dyker upp för det första respektive för den andra respektive för den tredje personen. Vi söker $P(X_1 = X_2 = X_3)$. Vi har:

$$P(X_1 = X_2 = X_3) = \sum_{k=1}^{\infty} P(\{X_1 = k\} \cap \{X_2 = k\} \cap \{X_3 = k\}) = |oberoendet| =$$

$$= \sum_{k=1}^{\infty} P(X_1 = k)P(X_2 = k)P(X_3 = k) =$$

$$= \sum_{k=1}^{\infty} \left((1-p)^{k-1}p\right)^3 = p^3 \sum_{k=1}^{\infty}((1-p)^3)^{k-1} = |\text{Ledning, ty } 0 < (1-1/6)^3 < 1| =$$

$$= p^3 \frac{1}{1-(1-p)^3} = \frac{1}{(\frac{1}{p})^3 - (\frac{1}{p}-1)^3} = |p = 1/6| = \frac{1}{6^3 - 5^3} = \frac{1}{91} \simeq 0.011$$

**Uppgift 15**

Först ska vi ta fram ML-skattningen av parametern $\beta$. Vi har

$$L(\beta) = \prod_{j=1}^{n} \frac{x_j^3 e^{-x_j/\beta}}{\beta^4 3!} = \left( (3!)^{-n} \prod_{j=1}^{n} x_j^3 \right) \beta^{-4n} e^{-n\bar{x}/\beta}$$

eller

$$\ln L(\beta) = \ln\left( (3!)^{-n} \prod_{j=1}^{n} x_j^3 \right) - 4n \ln\beta - n\bar{x}/\beta.$$

Derivering ger

$$\frac{d \ln L(\beta)}{d\beta} = -\frac{4n}{\beta} + \frac{n\bar{x}}{\beta^2}.$$

Av $\frac{d \ln L(\beta)}{d\beta} = 0$ fås

$$\frac{4n}{\beta} = \frac{n\bar{x}}{\beta^2} \Rightarrow \beta^*_{obs} = \frac{\bar{x}}{4}.$$

Eftersom $\ln L(\beta) \to -\infty$ då $\beta \to 0$ (eller $\infty$) så följer att extremvärdet är ett maximum och att $\beta^*_{obs} = \bar{x}/4$ är ML-skattningen av $\beta$.

Nu ska vi bestämma medelfelet $d(\beta^*)$ för $\beta^*$. Vi har för $n$ observationer:

$$V(\beta^*) = V(\bar{X}/4) = \frac{1}{16} \cdot \frac{1}{n^2} \cdot n\, V(X_1) = \frac{V(X_1)}{16n},$$

där

$$V(X_1) = E\big(X_1^2\big) - \big(E(X_1)\big)^2 = |\text{enl lydelse}| = 20\beta^2 - (4\beta)^2 = 4\beta^2.$$

Vi har därmed

$$V(\beta^*) = \frac{V(X_1)}{16n} = \frac{4\beta^2}{16n} = \frac{\beta^2}{4n}.$$

Därför,

$$D(\beta^*) = \sqrt{V(\beta^*)} = \sqrt{\frac{\beta^2}{4n}} = \frac{\beta}{2\sqrt{n}}\,.$$

Slutligen, medelfelet $d(\beta^*)$ för $\beta^*$ är som följer:

$$d(\beta^*) = \Big(D(\beta^*)\Big)^*_{obs} = \left(\frac{\beta}{2\sqrt{n}}\right)^*_{obs} = \frac{\beta^*_{obs}}{2\sqrt{n}} = \left|\beta^*_{obs} = \bar{x}/4\right| = \frac{\bar{x}/4}{2\sqrt{n}} = \frac{\bar{x}}{8\sqrt{n}} =$$

$$= \left|x_1 = 4,\ x_2 = 5,\ x_3 = 1,\ x_4 = 10\right| = \frac{20/4}{8\sqrt{4}} = \frac{5}{16} \simeq 0.3125$$

## Uppgift 16

Först beräknar vi felmarginalen. Vi har:

$$\text{felmarginal } = \frac{0.256 - (-5.076)}{2} = 2.666$$

Vi skriver om våra noll- och mothypoteser (med tanke på vad som efterfrågas, dvs $\mu_A = \mu_B + 5$):
$H_0 : \mu_A - \mu_B = 0$
$H_1 : \mu_A - \mu_B = 5$

Testets styrka nu då $\mu_A = \mu_B + 5$ är som följer:

$$h(5) = P(\text{förkasta } H_0 \,\big|\, H_1 \text{ är sann}) = P\big(0 \notin I_{\mu_A - \mu_B}(0.95) \,\big|\, \mu_A - \mu_B = 5\big) =$$

$$= 1 - P\big(0 \in I_{\mu_A - \mu_B}(0.95) \,\big|\, \mu_A - \mu_B = 5\big) =$$

$$= 1 - P\big(|\bar{X}_A - \bar{X}_B - 0| < \text{ felmarginal } \,\big|\, \mu_A - \mu_B = 5\big) =$$

$$= 1 - P\big(|\bar{X}_A - \bar{X}_B| < 2.666 \,\big|\, \mu_A - \mu_B = 5\big) =$$

$$= 1 - P\left(|\bar{X}_A - \bar{X}_B| < 2.666 \,\middle|\, \bar{X}_A - \bar{X}_B \in N\left(5, \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}\right)\right) =$$

$$= 1 - P\left(|\bar{X}_A - \bar{X}_B| < 2.666 \,\middle|\, \bar{X}_A - \bar{X}_B \in N\left(5, \frac{\text{felmarginal}}{\lambda_{\alpha/2}}\right)\right) =$$

$$= \big|\alpha = 0.05, \text{ Tabell 2}\big| = 1 - P\left(|\bar{X}_A - \bar{X}_B| < 2.666 \,\middle|\, \bar{X}_A - \bar{X}_B \in N\left(5, \frac{2.666}{1.96}\right)\right) =$$

$$= 1 - P\left(-2.666 < \bar{X}_A - \bar{X}_B < 2.666 \,\middle|\, \bar{X}_A - \bar{X}_B \in N\left(5, \frac{2.666}{1.96}\right)\right) =$$

$$= \big|\text{Sats 6.1}\big| = 1 - \left(\Phi\left(\frac{2.666 - 5}{\frac{2.666}{1.96}}\right) - \Phi\left(\frac{-2.666 - 5}{\frac{2.666}{1.96}}\right)\right) =$$

$$= 1 - \big(\Phi(-1.72) - \Phi(-5.64)\big) \simeq$$

$$\simeq \big|\text{Tabell 1}\big| \simeq 1 - \Phi(-1.72) + 0.0 =$$

$$= \Phi(1.72) = \big|\text{Tabell 1}\big| = 0.9573 \simeq 0.957$$

Dvs, testets styrka då $\mu_A = \mu_B + 5$ är ca 95.7% .