

Tentamentsskrivning i **Matematisk Statistik MVE091/MVE090**

Tid: den 9 april, 2026, 08:30-13:30

Examinator och jour: Erik Broman, mob. 073 7320791,

Hjälpmedel: Typgodkänd miniräknare, 4 A4-sidor egenhändigt skrivna anteckningar (2 ark fram och bak eller 4 ark på en sida) samt utdelade tabeller.

Tentamen består av 8 frågor om sammanlagt 50 poäng. Preliminära betygsgränser är satta till:

betyg "3": 20 till 29 poäng

betyg "4": 30 till 39 poäng

betyg "5": 40 eller fler poäng.

OBS! Alla lösningar skall vara väl redovisade, motiverade och fullständiga. Svar skall ges på enklast möjliga form. Talen är ej ordnade efter svårighetsgrad.

Följande formler kan eventuellt vara till nytta:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}, \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \sum_{k=0}^n x^k = \frac{1-x^{n+1}}{1-x}, \quad e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$$\lim_{n \rightarrow \infty} (1+a_n)^{b_n} = e^c \text{ om } \lim_{n \rightarrow \infty} a_n b_n = c \text{ och } \lim_{n \rightarrow \infty} a_n = 0.$$

1. Alice har två skålar med pärlor framför sig. I skål 1 finns det 2 vita, 2 röda och 3 blå pärlor medans det i skål två finns det 3 vita, 1 röd och 2 blå pärlor.

- (a) Antag att Alice först väljer en skål slumpmässigt. Sedan väljer Alice slumpmässigt två pärlor från den tidigare valda skålen. Vad är sannolikheten att båda pärlorna är vita? (2p)
- (b) Antag istället att Alice först väljer en skål slumpmässigt och tar en pärla (slumpmässigt) från den valda skålen. Sedan väljer Alice igen en skål slumpmässigt och tar en pärla (slumpmässigt) från den valda skålen. Antag att alla val är oberoende från varandra. Vad är nu sannolikheten att båda pärlorna är vita? Alice lägger inte tillbaka den första pärlan innan hon tar den andra. (2p)
- (c) Antag till slut att Alice tröttnar på att välja skålar och samlar ihop alla pärlorna i en enda skål. Sedan väljer hon två pärlor slumpmässigt ur den enda skålen. Vad är nu sannolikheten att pärlorna har samma färg? (2p)

Lösning:

- (a) Låt A_k vara händelsen att Alice valda skål nummer k för $k = 1, 2$. Låt sedan VV vara händelsen att Alice valde två vita pärlor. Vi söker då

$$\mathbb{P}(VV) = \sum_{k=1}^2 \mathbb{P}(VV|A_k) \mathbb{P}(A_k) = \frac{1}{2} \left(\frac{\binom{2}{2}}{\binom{7}{2}} + \frac{\binom{3}{2}}{\binom{6}{2}} \right) = \frac{13}{105} (\approx 0.124).$$

- (b) Låt A_k vara händelsen att Alice valda skål nummer k för $k = 1, 2$ i det första valet, och låt B_k vara händelsen att Alice valda skål nummer k för $k = 1, 2$ i det andra valet. Vi har nu att

$$\begin{aligned} \mathbb{P}(VV) &= \frac{1}{4} (\mathbb{P}(VV|A_1 \cap B_1) + \mathbb{P}(VV|A_1 \cap B_2) + \mathbb{P}(VV|A_2 \cap B_1) + \mathbb{P}(VV|A_2 \cap B_2)) \\ &= \frac{1}{4} \left(\frac{\binom{2}{2}}{\binom{7}{2}} + 2 \frac{2}{7} \frac{3}{6} + \frac{\binom{3}{2}}{\binom{6}{2}} \right) = \frac{2}{15} (\approx 0.133). \end{aligned}$$

- (c) Låt VV, RR, BB vara händelserna att Alica valde två vita, två röda respektive två blå pärlor. Vi söker då

$$\mathbb{P}(VV) + \mathbb{P}(RR) + \mathbb{P}(BB) = \frac{\binom{5}{2}}{\binom{13}{2}} + \frac{\binom{3}{2}}{\binom{13}{2}} + \frac{\binom{5}{2}}{\binom{13}{2}} = \frac{23}{78} (\approx 0.295)$$

2. Jöns-Harald färdas tillbaka i tiden och till ett parallellt universum. Väl där utses han till astronaut och placeras helt sonika på en raket med destination månen. I detta parallella universum har transistorn ännu inte uppfunnits och därför måste alla beräkningar ske med hjälp av en dator som är konstruerad av 700 vakuumrör.

Jöns-Haralds mändator är byggd med viss redundans och fungerar ifall minst 550 av vakuumrören fungerar. Livslängderna $T_1, \dots, T_{700}$ (timmar) för vakuumrören kan anses vara oberoende och följa en Weibull-fördelning med parametrar α, β , dvs

$$f_T(s) = \alpha \beta s^{\beta-1} e^{-\alpha s^\beta} \text{ för } s \geq 0,$$

där det är känt att $\alpha = 1/40$ och $\beta = 1/2$.

- (a) Bestäm sannolikheten att ett specifikt vakuumrör fungerar efter exakt fyra dygn. (1p)
- (b) Om datorn inte fungerar så kan raketten inte återvända och Jöns-Harald dör. Beräkna sannolikheten att Jöns-Harald dör som en följd av datorfel om månturen tar exakt fyra dygn. (3p)
- (c) Jöns-Harald blir lite försenad hem från månen och därför ökar chansen att vakuumrören hinner gå sönder. Beräkna (med hygglig approximation) hur stor chansen måste vara för att ett vakuumrör är sönder om risken att Jöns-Harald dör av datorfel skall vara 90%? (3p)

Lösning:

- (a) Sannolikheten ges av

$$\mathbb{P}(T \geq 96) = \int_{96}^{\infty} \alpha \beta s^{\beta-1} e^{-\alpha s^\beta} ds = \left[-e^{-\alpha s^\beta} \right]_{96}^{\infty} = e^{-\alpha 96^\beta} = e^{-\sqrt{96}/40} (\approx 0.7825).$$

- (b) Låt X vara antalet vakuumrör som fungerar efter exakt fyra dygn. Vi ser att

$$X \sim \text{Bin}(n, p) \text{ där } n = 700 \text{ och } p = e^{-\sqrt{96}/40}.$$

Vi söker $\mathbb{P}(X < 550)$ och här är det lämpligt med normalapproximation. Vi får att $X \approx N(np, np(1-p))$ så att

$$\mathbb{P}(X < 550) = \mathbb{P}\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{550 - np}{\sqrt{np(1-p)}}\right) \approx \mathbb{P}(Z < 0.19) \approx 0.576.$$

- (c) Vi har att $\mathbb{P}(Z \leq -1.28) \approx 0.1$ så vi måste lösa ekvationen

$$\frac{550 - np}{\sqrt{np(1-p)}} = -1.28$$

med avseende på p . Man kan ju såklart försöka lösa denna med avseende på p , men det är mycket enklare att försöka testa sig fram. Om vi tar $p = 0.8$ så får vi svaret ≈ -0.95 medans 0.81 ger svaret ≈ -1.64 och letar man ett tag till så ser man att $p = 0.805$ ger svaret ≈ -1.288 som duger fint.

3. Låt X vara likformigt fördelad på mängden $\{-1, 0, 1, 2\}$.

(a) Låt $Y = X^2$. Bestäm täthetsfunktionen för Y . (2p)

(b) Bestäm de momengenererande funktionerna för X och för Y . (2p)

(c) Låt $Z = Y_1 + Y_2$ där Y_1, Y_2 är två oberoende slumpvariabler med samma täthetsfunktion som Y . Beräkna täthetsfunktionen för Z . (2p)

Lösning:

(a) Vi ser att $Y \in \{0, 1, 4\}$ och har att

$$\mathbb{P}(Y = 0) = \mathbb{P}(X = 0) = \frac{1}{4}, \quad \mathbb{P}(Y = 4) = \mathbb{P}(X = 2) = \frac{1}{4}$$

och

$$\mathbb{P}(Y = 1) = \mathbb{P}(X = -1) + \mathbb{P}(X = 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

(b) Vi får att

$$M_X(t) = \mathbb{E}[e^{Xt}] = \sum_{k=-1}^2 e^{kt} \mathbb{P}(X = k) = \frac{e^{-t} + e^{0t} + e^t + e^{2t}}{4} = \frac{e^{-t} + 1 + e^t + e^{2t}}{4}$$

och att

$$M_Y(t) = \mathbb{E}[e^{Yt}] = \frac{1 + 2e^t + e^{4t}}{4}.$$

(c) Då $Y_1, Y_2 \in \{0, 1, 4\}$ ser vi att $Z \in \{0, 1, 2, 4, 5, 8\}$ med

$$\mathbb{P}(Z = 0) = \mathbb{P}(Y_1 = Y_2 = 0) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$$

$$\mathbb{P}(Z = 1) = 2\mathbb{P}(Y_1 = 0, Y_2 = 1) = 2 \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{4}$$

$$\mathbb{P}(Z = 2) = \mathbb{P}(Y_1 = Y_2 = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\mathbb{P}(Z = 4) = 2\mathbb{P}(Y_1 = 0, Y_2 = 4) = 2 \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{8}$$

$$\mathbb{P}(Z = 5) = 2\mathbb{P}(Y_1 = 1, Y_2 = 4) = 2 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\mathbb{P}(Z = 8) = \mathbb{P}(Y_1 = Y_2 = 4) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}.$$

4. Låt X vara exponentialfördelad med parameter $\beta > 0$, så att X har täthetsfunktion

$$f_X(x) = \frac{1}{\beta} e^{-x/\beta} \text{ för } x > 0.$$

Givet att $X = x$, låt sedan Y vara exponentialfördelad med parameter $1/x$.

(a) Bestäm den gemensamma täthetsfunktionen för X och Y . (2p)

(b) Bestäm marginaltäthetsfunktionen för Y . (2p)

(c) Beräkna $\mathbb{E}[XY]$. (2p)

Lösning:

(a) Vi har att

$$f_{Y|x}(y) = xe^{-yx} \text{ för } y > 0,$$

så att

$$f_{X,Y}(x,y) = f_{Y|x}(y)f_X(x) = xe^{-yx}\frac{1}{\beta}e^{-x/\beta} \text{ för } x,y > 0.$$

(b) Vi har att

$$\begin{aligned} f_Y(y) &= \int_0^\infty xe^{-yx}\frac{1}{\beta}e^{-x/\beta}dx = \frac{1}{\beta} \int_0^\infty xe^{-x(y+1/\beta)}dx \\ &= \frac{1}{\beta(y+1/\beta)} \int_0^\infty x(y+1/\beta)e^{-x(y+1/\beta)}dx = \frac{1}{\beta(y+1/\beta)^2} \text{ för } y > 0, \end{aligned}$$

där den sista likheten kommer sig av att integralen är väntevärdet av en exponentialfördelad slumpvariabel med parameter $1/(y+1/\beta)$.

(c) Vi har att

$$\begin{aligned} \mathbb{E}[XY] &= \int_{-\infty}^\infty \int_{-\infty}^\infty xyf_{X,Y}(x,y)dxdy = \int_0^\infty \int_0^\infty xyxe^{-yx}\frac{1}{\beta}e^{-x/\beta}dxdy \\ &= \int_0^\infty x\frac{1}{\beta}e^{-x/\beta} \int_0^\infty xye^{-yx}dydx = \int_0^\infty x\frac{1}{\beta}e^{-x/\beta}\frac{1}{x}dx = \int_0^\infty \frac{1}{\beta}e^{-x/\beta}dx = 1, \end{aligned}$$

eftersom $\int_0^\infty xye^{-yx}dy$ är väntevärdet av en exponentialfördelad slumpvariabel med parameter $1/x$.

5. Ett system består av tre stycken seriekopplade komponenter. Livslängden (i timmar) för komponenterna anses vara oberoende och Gamma-fördelade med paramterar $\alpha = 3$ och $\beta = 1$ så att täthetsfunktionen blir

$$\frac{1}{\Gamma(3)}x^2e^{-x} = \frac{1}{2}x^2e^{-x} \text{ för } x \geq 0.$$

Systemet fallerar (dvs går sönder) ifall någon av komponenterna fallerar.

- (a) Beräkna sannolikheten att systemet fallerar inom två timmar. (4p)
 (b) Beräkna sannolikheten att systemet fallerar inom två timmar om du vet att den tredje komponenten fungerar efter 2 timmar. (1p)

Lösning:

- (a) Låt T_1, T_2, T_3 beteckna livslängderna för de tre komponenterna. Vi söker

$$\begin{aligned} \mathbb{P}(\min(T_1, T_2, T_3) \leq 2) &= 1 - \mathbb{P}(\min(T_1, T_2, T_3) \geq 2) \\ &= 1 - \mathbb{P}(T_1 \geq 2)\mathbb{P}(T_2 \geq 2)\mathbb{P}(T_3 \geq 2) = 1 - \mathbb{P}(T_1 \geq 2)^3, \end{aligned}$$

där

$$\begin{aligned} \mathbb{P}(T_1 \geq 2) &= \int_2^\infty \frac{1}{2}x^2e^{-x}dx = \frac{1}{2}[-x^2e^{-x}]_2^\infty + \int_2^\infty xe^{-x}dx \\ &= 2e^{-2} + [-xe^{-x}]_2^\infty + \int_2^\infty e^{-x}dx = 4e^{-2} + [-e^{-x}]_2^\infty = 5e^{-2} \end{aligned}$$

så svaret blir

$$\mathbb{P}(\min(T_1, T_2, T_3) \leq 2) = 1 - (5e^{-2})^3.$$

(b) Vi söker här

$$\mathbb{P}(\min(T_1, T_2, T_3) \leq 2 | T_3 > 2) = \mathbb{P}(\min(T_1, T_2) \leq 2) = 1 - \mathbb{P}(T_1 \geq 2)^2 = 1 - (5e^{-2})^2.$$

6. Lös följande problem.

(a) Låt X vara en slumpvariabel med täthetsfunktion

$$f_X(x) = \frac{6x}{\theta^2} \left(1 - \frac{x}{\theta}\right) \text{ för } 0 \leq x \leq \theta.$$

Hitta momentskattaren (MME'n) för θ . (3p)

(b) Är ditt svar i (a) en väntevärdesriktig skattare för θ ? (1p)

(c) Låt $X_1, \dots, X_n$ vara en sekvens av oberoende och likafördelade diskreta slumpvariabler sådana att

$$\mathbb{P}(X_k = l) = \begin{cases} \theta & \text{om } l = 0 \\ \frac{1-\theta}{3} & \text{om } l = 1 \\ \frac{2(1-\theta)}{3} & \text{om } l = 2. \end{cases}$$

Bestäm likelihooden för θ uttryckt i termer av n_0, n_1, n_2 där

$$n_l = \#\{k : X_k = l\} \text{ (i ord: antalet } X_k \text{ sådana att } X_k = l)$$

för $l = 0, 1, 2$. Här är parametern $0 < \theta < 1$. (1p)

(d) Hitta MLE:n för θ baserat på ditt svar i (c). (3p)

Lösning:

(a) Vi har att

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^{\theta} \frac{6x^2}{\theta^2} \left(1 - \frac{x}{\theta}\right) dx \\ &= \frac{6}{\theta^2} \int_0^{\theta} x^2 dx - \frac{6}{\theta^3} \int_0^{\theta} x^3 dx = \frac{6}{\theta^2} \frac{\theta^3}{3} - \frac{6}{\theta^3} \frac{\theta^4}{4} = 2\theta - \frac{3\theta}{2} = \frac{\theta}{2}. \end{aligned}$$

Momentmetoden ger då att $\hat{\theta} = 2\bar{X}$.

(b) Vi ser att

$$\mathbb{E}[\hat{\theta}] = \mathbb{E}[2\bar{X}] = 2\mathbb{E}\left[\frac{1}{n} \sum_{k=1}^n X_k\right] = \frac{2}{n} \sum_{k=1}^n \mathbb{E}[X_k] = \frac{2}{n} \sum_{k=1}^n \frac{\theta}{2} = \frac{2}{n} n \frac{\theta}{2} = \theta.$$

(c) Vi har att

$$L(\theta) = \prod_{k=1}^n f(X_k|\theta) = \theta^{n_0} \left(\frac{1-\theta}{3}\right)^{n_1} \left(\frac{2(1-\theta)}{3}\right)^{n_2}.$$

(d) Vi får att

$$l(\theta) = n_0 \log(\theta) + n_1 (\log(1-\theta) - \log(3)) + n_2 (\log(2) + \log(1-\theta) - \log(3))$$

och därmed ser vi att

$$l'(\theta) = \frac{n_0}{\theta} - \frac{n_1 + n_2}{1-\theta}$$

och att

$$l''(\theta) = -\frac{n_0}{\theta^2} - \frac{n_1 + n_2}{(1-\theta)^2} < 0,$$

så att $l'(\theta) = 0$ kommer ge oss ett maximum. Vi ser att

$$l'(\theta) = 0 \iff n_0(1 - \theta) = (n_1 + n_2)\theta \iff n_0 = (n_0 + n_1 + n_2)\theta = n\theta$$

så vår MLE blir

$$\hat{\theta} = \frac{n_0}{n}.$$

7. Betrakta ett stickprov $X_1, \dots, X_n$ av likafördelade och oberoende slumpvariabler som är Poisson-fördelade med parameter $\lambda > 0$.

- (a) Härled ett uttryck för ett $100(1 - \alpha)\%$ två-sidigt konfidensintervall för λ . Man kan här anta att n är stort. (4p)
- (b) Martinas chef har gjort några experiment och ger Martina följande dataserie

Experiment nummer:	1	2	3	4	5
Resultat:	2	0	1	1	4.

Chefen påstår att dataserien kommer från en Poissonfördelning. Använd ditt svar i (a) för att ange ett numeriskt 99% konfidensintervall för λ . Kommentera ditt svar. (2p)

Lösning:

- (a) Vi har enligt centrala gränsvärdessatsen att

$$\bar{X} \approx N(\mathbb{E}[\bar{X}], \text{Var}(\bar{X})) = N\left(\lambda, \frac{\lambda}{n}\right)$$

eftersom

$$\mathbb{E}[\bar{X}] = \mathbb{E}[X_1] = \lambda \text{ och } \text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = \frac{\lambda}{n}$$

där vi använde våra kunskaper om Poisson-fördelningen. Vi ser att en lämplig referensvariabel blir

$$\frac{\bar{X} - \lambda}{\sqrt{\lambda/n}} \approx \frac{\bar{X} - \lambda}{\sqrt{\bar{X}/n}} \approx N(0, 1),$$

där vi använde "ingenjörstricket" i den första approximationen. Vi ser att

$$\begin{aligned} 1 - \alpha &= \mathbb{P}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) \approx \mathbb{P}\left(-z_{\alpha/2} \leq \frac{\bar{X} - \lambda}{\sqrt{\bar{X}/n}} \leq z_{\alpha/2}\right) \\ &= \mathbb{P}\left(-z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}} \leq \bar{X} - \lambda \leq z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}}\right) = \mathbb{P}\left(\bar{X} - z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}} \leq \lambda \leq \bar{X} + z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}}\right), \end{aligned}$$

där $Z \sim N(0, 1)$. Det sökta konfidensintervallet blir därför

$$I_\lambda = \bar{X} \pm z_{\alpha/2} \sqrt{\frac{\bar{X}}{n}}.$$

- (b) Vi har att $z_{0.005} \approx 2.575$ och $\bar{x} = \frac{8}{5}$. Därmed blir ett numeriskt 99% konfidensintervall för λ

$$I_\lambda = \bar{x} \pm z_{0.005} \sqrt{\frac{\bar{x}}{5}} \approx [0.14, 3.06].$$

Martina bör informera sin chef om att uttrycket är framtaget för stora stickprov och bör inte appliceras för så små stickprov som $n = 5$. Svaret riskerar att vara grovt missvisande. Martina vet inte heller ifall chefen har gjort sina experiment oberoende av varandra och bör fråga om detta.

8. Ett forskarteam misstänker att en viss, relativt vanlig genvariant kan leda till minskad tillväxt av fotlängden hos bärare av genvarianten. För att undersöka detta mätte forskarna fotlängden hos två grupper av vuxna kvinnor. I en grupp bar alla på genvarianten och i den andra var det ingen som bar på genvarianten. Båda grupperna bestod av 30 individer. Resultaten sammanfattas som följer:

Grupp:	Med genvariant	Utan genvariant
Genomsnittlig fotlängd (i cm):	23.3	24.6
Skattad varians (s^2):	1.4	1.8.

Forskarteamet utgår ifrån att de uppmätta längderna kommer från normalfördelningar med samma varians. De antar dessutom att resultaten av alla mätningar är oberoende av varandra.

- Sätt upp ett lämpligt hypotestest för att testa forskarteamets hypotes. (1p)
- Ange en lämplig test-statistika för att genomföra hypotestestet i del (a). Förklara de ingående delarna och ange fördelningen för test-statistikan. (2p)
- Genför testet på signifikansnivån $\alpha = 0.05$ och ange p-värdet. Skall du förkasta H_0 eller ej? (3p)

Lösning:

- Låt μ_u respektive μ_m vara snittlängden av fötterna i gruppen utan gen respektive med gen. Forskarna vill testa

$$\begin{aligned} H_0 : & \mu_u = \mu_m \\ H_1 : & \mu_u < \mu_m. \end{aligned}$$

- Låt $X = (X_1, \dots, X_{30})$ beteckna stickprovet från gruppen med genvarianten och låt $Y = (Y_1, \dots, Y_{30})$ beteckna stickprovet från gruppen utan genvarianten. Vi har att

$$\bar{X} - \bar{Y} \sim N\left(\mu_u - \mu_m, \frac{\sigma^2}{30} + \frac{\sigma^2}{30}\right)$$

där σ^2 är variansen. En lämplig test-statistika blir därför

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{s_p^2/15}} \text{ där } s_p^2 = \frac{29s_u^2 + 29s_m^2}{58} = 1.6,$$

och där s_u^2 respektive s_m^2 är skattningarna av variansen i de två stickproven. Vi har att $T \stackrel{H_0}{\sim} t(58)$, men om H_0 inte är sant så har T en annan fördelning.

- Värdet på test-statistikan blir i detta fall

$$T(x, y) = \frac{\bar{x} - \bar{y}}{\sqrt{s_p^2/15}} \approx -3.98.$$

Vi ser också att om H_1 skulle vara sant, så borde $\bar{X} - \bar{Y} < 0$ så vi förkastar H_0 till förmån för H_1 om vi får ett stort, negativt värde på T . Vi ser att

$$\text{p-värde} = \mathbb{P}(T(X, Y) \leq T(x, y)) = \mathbb{P}(T(X, Y) \leq -3.98) \leq \mathbb{P}(T(X, Y) \leq -3.466) = 0.0005,$$

där vi använde det mest extrema värde vi kunde hitta i tabellen. p-värdet är mycket litet och vi förkastar därför H_0 till förmån för H_1 .