Tentamentsskrivning i **Matematisk Statistik TMA321**

Tid: den 12 Oktober, 2019
Hjälpmedel: Typgodkänd miniräknare, egenhändigt skriven formelsamling om två A4 fram och bak (dvs 4 sidor), 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 och motiverade. Talen är ej ordnade efter svårighetsgrad.

1. Låt X_1, X_2 vara oberoende slumpvariabler som båda är Binomialfördelade med parametrar n, p. Betrakta de tre skattarna

$$\hat{p} = \frac{X_1}{n}, \quad p^* = \frac{X_1 + 1}{n+2} \quad \text{och } \tilde{p} = \frac{X_1^2 - X_1 X_2}{n} + \frac{X_1 X_2}{n^2}.$$

   (a) Vilken/vilka av de tre skattarna är väntevärdesriktiga? (3p)

   (b) Under antagandet att $n = 4$ och $p = 1/4$, vilken av skattarna $\hat{p}$ och p^* har minsta förväntade kvadratiska avvikelsen från p? (dvs vilken av $\mathbb{E}[(\hat{p} - p)^2]$ och $\mathbb{E}[(p^* - p)^2]$ är minst?) (3p)

   (c) Varför är svaret i (b) anmärkningsvärt? (1p)

2. Saga älskar frukostflingorna ”Späsial Q” för varje paket innehåller alltid exakt en färgglad leksak. Det finns 5 olika leksaker och varje leksak har samma sannolikhet att hamna i ett givet paket. Dessutom kan vi anta att vilken leksak som hamnar i ett visst paket är oberoende mellan paketen.

   Saga vill samla alla fem olika leksaker.

   (a) Antag att Saga hittills har samlat tre unika leksaker. Låt X vara antalet paket av ”Späsial Q” som hon måste köpa till och med att hon hittar en leksak som hon inte redan har. Vilken sannolikhetsfunktion har X? Vilket väntevärde har X? (3p)

   (b) Vilket är det förväntade antalet paket som Saga måste köpa till och med att hon samlat ihop alla fem leksakerna? (Här skall även paketen som behövs för att Saga skall samla de första tre räknas in.) (3p)

3. Låt U vara likformigt fördelad på intervallet $[0, 1]$ (dvs $U \sim$U$([0, 1])$), och givet att $U = u$, låt T vara exponentialfördelad med parameter $1/u$ (dvs $T \sim$Exp$(1/U)$).

   (a) Beräkna den momentgenererande funktionen för T. (2p)

(b) Beräkna den momentgenererande funktionen för slumpvariabeln $V = \log U$. (2p)

(c) Använd svaret i (b) för att beräkna

$$\mathbb{E}[(\log U)^k],$$

för $k = 1, 2, \ldots$ (2p)

4. Simon gillar en konstig variant av poker där alla deltagarna får 3 kort som är dragna från en standardkortlek bestående av 52 kort.

(a) Hur många 3-korts händer innehåller minst två kort av samma valör? (3p)

(b) Vad är sannolikheten att en 3-korts hand innehåller en triss, givet att den innehåller minst två av samma valör? (3p)

5. Ottilia gillar att experimentera med trolldeg. Hon blandar i olika mängd av ämnet V till sin mix och testar hur detta påverkar elasticiteten. Hon rullar degen till en boll som hon sedan släpper från ett bord för att se hur högt den studsar tillbaka.

Hon upprepar experimentet 7 gånger. I tabellen nedan ser vi resultatet av Ottilias experiment.

Mängd V (i gram):	1	2	3	4.5	6	8	10
Studshöjd (i cm):	17.8	21.4	23.6	24	25.8	27	27.3

Ottilia ansätter en linjär regressionsmodell $y = \beta_0 + \beta_1 x$ där y är studshöjden och x är mängden av ämnet V. Data sammanfattas med att $S_{xx} \approx 64.2$, $S_{yy} \approx 68.3$ och $S_{xy} = 60.6$.

(a) Skatta β_0 och β_1. (2p)

(b) Ottilia misstänker att studshöjden ökar precis i takt med mängden V (dvs att $\beta_1 = 1$). Understöder data hennes hypotes? (2p)

(c) Ange förklaringsgraden och beräkna residualerna. Kommentera ditt resultat. (2p)

6. Låt X vara en slumpvariabel med täthetsfunktion

$$f_X(x) = C_\alpha x^\alpha, \text{ för } 0 \leq x \leq 1,$$

där $\alpha > 0$ är en parameter.

(a) För varje fixt värde på α, bestäm C_α så att $f_X(x)$ verkligen utgör en täthetsfunktion. (1p)

(b) Låt $Y = X^2 + 1$, bestäm täthetsfunktionen för Y. (3p)

(c) Bestäm täthetsfunktionen för $W = \sqrt{Y}$. (3p)

7. Egil gillar guld och är därför guldgrävare. Han gräver fram guldhaltigt grus som krossas till ett fint damm ur vilket han sedan extraherar själva guldet. Egil använder sig av en traditionell smältmetod, men funderar på att byta till en kemiskt baserad metod (han dränker helt sonika allt damm i cyanid).

   Då den nya metoden är farlig och relativt dyr så vill han veta om den faktiskt ger bättre resultat än den gamla. Han tar därför fram åtta stendammprover från åtta olika markplättar. Varje prov delas sedan upp i två delar och metoderna jämförs.

Prov nr:	1	2	3	4	5	6	7	8
Halt (metod 1, ppm):	6.5	8.1	7.2	4.1	5.9	2.6	7.7	1.6
Halt (metod 2, ppm):	7.4	8.9	8.5	5.4	5.6	3.6	8.5	3.2

   Egil antar att proverna kommer från normalfördelningar.

   (a) Sätt upp lämpligt hypotestest för att testa om det är någon skillnad (positiv eller negativ) på signifikansnivån 1 %. (3p)

   (b) För att det skall vara lönsamt att byta metod måste skillnaden vara större än 0.3 ppm. Sätt upp ett lämpligt hypotestest för att testa huruvida så är fallet. Använd samma signifikansnivå som i uppgift (a). (3p)

8. Låt X ha täthetsfunktion

$$f_X(x) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2} \text{ för } -\infty < x < \infty,$$

   där $\alpha > 0$ är okänd.

   (a) Hitta momentskattaren (MME:n) för α. (3p)

   (b) Hitta maximum likelihood skattaren (MLE:n) för α. (3p)

Tentamentsskrivning i **Matematisk Statistik TMA321**

Tid: den 12 Oktober, 2019
Hjälpmedel: Typgodkänd miniräknare, egenhändigt skriven formelsamling om två A4 fram och bak (dvs 4 sidor), 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 och motiverade. Talen är ej ordnade efter svårighetsgrad.

1. Låt X_1, X_2 vara oberoende slumpvariabler som båda är Binomialfördelade med parametrar n, p. Betrakta de tre skattarna

$$\hat{p} = \frac{X_1}{n}, \quad p^* = \frac{X_1 + 1}{n+2} \quad \text{och } \tilde{p} = \frac{X_1^2 - X_1 X_2}{n} + \frac{X_1 X_2}{n^2}.$$

(a) Vilken/vilka av de tre skattarna är väntevärdesriktiga? (3p)

(b) Under antagandet att $n = 4$ och $p = 1/4$, vilken av skattarna $\hat{p}$ och p^* har minsta förväntade kvadratiska avvikelsen från p? (dvs vilken av $\mathbb{E}[(\hat{p} - p)^2]$ och $\mathbb{E}[(p^* - p)^2]$ är minst?) (3p)

(c) Varför är svaret i (b) anmärkningsvärt? (1p)

Lösning:

(a) Vi har att

$$\mathbb{E}[\hat{p}] = \frac{\mathbb{E}[X_1]}{n} = \frac{np}{n} = p,$$
$$\mathbb{E}[p^*] = \frac{\mathbb{E}[X_1] + 1}{n+2} = \frac{np+1}{n+2} \neq p$$

och vidare har vi att

$$\begin{aligned}
\mathbb{E}[\tilde{p}] &= \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1 X_2]}{n} + \frac{\mathbb{E}[X_1 X_2]}{n^2} \\
&= \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{n} + \frac{\mathbb{E}[X_1]^2}{n^2} = \frac{\mathrm{Var}(X_1)}{n} + \frac{\mathbb{E}[X_1]^2}{n^2} \\
&= \frac{np(1-p)}{n} + \frac{(np)^2}{n^2} = p(1-p) + p^2 = p.
\end{aligned}$$

Vi ser alltså att $\hat{p}$ och $\tilde{p}$ är väntevärdesriktiga.

(b) Vi har att

$$\mathbb{E}[(\hat{p}-p)^2] = \mathbb{E}\left[\left(\frac{X_1}{n}-p\right)^2\right] = \mathrm{Var}\left(\frac{X_1}{n}\right) = \frac{\mathrm{Var}(X_1)}{n^2} = \frac{p(1-p)}{n},$$

så att $\mathbb{E}\left[\left(\hat{p}-\frac{1}{4}\right)^2\right] = \frac{3}{64}$ då $n=4$. För p^* ser vi att

$$\begin{aligned}\mathbb{E}\left[(p^*-p)^2\right] &= \mathbb{E}\left[\left(\frac{X_1+1}{n+2}-p\right)^2\right] = \frac{\mathbb{E}[(X_1+1-np-2p)^2]}{(n+2)^2}\\ &= \frac{\mathbb{E}[(X_1-np)^2]+\mathbb{E}[(1-2p)^2]+\mathbb{E}[2(X_1-np)(1-2p)]}{(n+2)^2}\\ &= \frac{\mathrm{Var}(X_1)+(1-2p)^2+0}{(n+2)^2} = \frac{np(1-p)+(1-2p)^2}{(n+2)^2}\\ &= \{n=4, p=1/4\} = \frac{3/4+1/4}{36} = \frac{1}{36} < \frac{3}{64}.\end{aligned}$$

(c) Skattaren p^* kan tolkas som att man har $n+2$ datapunkter varav två stycken är fixerade till 1/2. Vi vet att denna fixering är felaktigt (då $p=1/4$), men trots detta kommer skattaren p^* ha minst kvadratisk avvikelse från det korrekta värdet 1/4.

2. Saga älskar frukostflingorna ”Späsial Q” för varje paket innehåller alltid exakt en färgglad leksak. Det finns 5 olika leksaker och varje leksak har samma sannolikhet att hamna i ett givet paket. Dessutom kan vi anta att vilken leksak som hamnar i ett visst paket är oberoende mellan paketen.

   Saga vill samla alla fem olika leksaker.

   (a) Antag att Saga hittills har samlat tre unika leksaker. Låt X vara antalet paket av ”Späsial Q” som hon måste köpa till och med att hon hittar en leksak som hon inte redan har. Vilken sannolikhetsfunktion har X? Vilket väntevärde har X? (3p)

   (b) Vilket är det förväntade antalet paket som Saga måste köpa till och med att hon samlat ihop alla fem leksakerna? (Här skall även paketen som behövs för att Saga skall samla de första tre räknas in.) (3p)

Lösning:

(a) Sannolikheten att nästa paket innehåller en av de två leksakerna som hon inte har är så klart 2/5. Då varje paket innehåller en ny leksak oberoende av varandra så ser vi att X är geometriskt fördelad. Vi har därför att $\mathbb{P}(X=k) = (3/5)^{k-1}2/5$ för $k=1,2,\ldots$. För en geometrisk fördelning gäller att $\mathbb{E}[X]=1/p$ och p är här 2/5. Vi ser därför att $\mathbb{E}[X] = 1/(2/5) = 5/2$.

(b) Låt nu X_k vara antalet paket som Saga måste köpa från det att hon har $k-1$ leksaker till och med att hon fått leksak nummer k. Vi ser då att X i uppgift (a) motsvarar X_4. Metoden i uppgift a ger direkt att $\mathbb{E}[X_1] = 1, \mathbb{E}[X_2] = 5/4, \mathbb{E}[X_3] = 5/3, \mathbb{E}[X_4] = 5/2$ och att $\mathbb{E}[X_5] = 5$. Den sökta siffran är $Y = X_1 + X_2 + X_3 + X_4 + X_5$ och vi ser att

$$\mathbb{E}[Y] = \sum_{k=1}^{5} \mathbb{E}[X_k] = 1 + 5/4 + 5/3 + 5/2 + 5 = 137/12.$$

3. Låt U vara likformigt fördelad på intervallet $[0,1]$ (dvs $U \sim$U$([0,1])$, och givet att $U = u$, låt T vara exponentialfördelad med parameter $1/u$ (dvs $T \sim$Exp$(1/U)$).

(a) Beräkna den momentgenererande funktionen för T. (2p)

(b) Beräkna den momentgenererande funktionen för slumpvariabeln $V = \log U$. (2p)

(c) Använd svaret i (b) för att beräkna

$$\mathbb{E}[(\log U)^k],$$

för $k = 1, 2, \ldots$ (2p)

Lösning:

(a) Om $Y \sim$Exp(λ) så vet vi att $M_Y(t) = \frac{\lambda}{\lambda - t}$. Därför blir

$$\mathbb{E}[e^{Ts}|U = u] = \frac{1/u}{1/u - s} = \frac{1}{1 - us},$$

så att $\mathbb{E}[e^{Ts}|U] = \frac{1}{1-Us}$. Det följer då att

$$\begin{aligned} M_T(s) &= \mathbb{E}[e^{Ts}] = \mathbb{E}\left[\mathbb{E}[e^{Ts}|U]\right] \\ &= \mathbb{E}\left[\frac{1}{1-Us}\right] = \int_0^1 \frac{1}{1-us} du = \left[\frac{\log(1-us)}{-s}\right]_0^1 = \frac{\log(1-s)}{-s} \end{aligned}$$

för $s < 1$.

(b) Vi har att

$$M_V(t) = \mathbb{E}[e^{Vt}] = \mathbb{E}[e^{t \log U}] = \mathbb{E}[U^t] = \int_0^1 u^t du = \left[\frac{u^{t+1}}{t+1}\right]_0^1 = \frac{1}{t+1},$$

för $t > -1$.

(c) Vi använder oss av att $\mathbb{E}[X^k] = M_X^{(k)}(t)|_{t=0}$ för $k = 1, 2, \ldots$. Vidare ser vi att

$$M_V'(t) = \frac{-1}{(t+1)^2}, \quad M_V''(t) = \frac{2}{(t+1)^3}, \quad M_V^{(k)}(t) = \frac{k!(-1)^k}{(t+1)^{(k+1)}}$$

så att

$$\mathbb{E}[(\log U)^k] = (-1)^k k!$$

för $k = 1, 2, \ldots$.

4. Simon gillar en konstig variant av poker där alla deltagarna får 3 kort som är dragna från en standardkortlek bestående av 52 kort.

   (a) Hur många 3-korts händer innehåller minst två kort av samma valör? (3p)

   (b) Vad är sannolikheten att en 3-korts hand innehåller en triss, givet att den innehåller minst två av samma valör? (3p)

Lösning:

(a) Det enklaste här är att lösa dela upp det i två subfall. I det första fallet är alla tre av samma valör. Vi ser att vi t.ex. kan få tre Äss på $\binom{4}{3}$ olika sätt och då vi har 13 olika valörer blir antalet händer med triss

$$13\binom{4}{3} = 52.$$

Vidare blir antalet händer med par i Äss $\binom{4}{2}$ som sedan kan kompletteras med ett tredje kort på 48 olika sätt. Alltså finns det $48\binom{4}{2}$ händer med par i Äss. Då man kan ha par i 13 olika valörer får vi då sammanlagt

$$13 \cdot 48 \cdot \binom{4}{2} = 3744,$$

händer med par. Det sökta antalet blir därför

$$13\binom{4}{3} + 13 \cdot 48 \cdot \binom{4}{2} = 3796.$$

(b) Låt T vara händelsen att handen innehåller en triss och låt A vara händelsen att handen innehåller minst två av samma valör. Vi efterfrågar då

$$\mathbb{P}(T|A) = \frac{\mathbb{P}(T \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(T)}{\mathbb{P}(A)}.$$

Vidare är

$$\mathbb{P}(T) = \frac{13\binom{4}{3}}{\binom{52}{3}},$$

och enligt svaret i (a) har vi att

$$\mathbb{P}(A) = \frac{13\binom{4}{3} + 13 \cdot 48 \cdot \binom{4}{2}}{\binom{52}{3}},$$

och därför blir

$$\mathbb{P}(T|A) = \frac{\mathbb{P}(T)}{\mathbb{P}(A)} = \frac{13\binom{4}{3}}{13\binom{4}{3} + 13 \cdot 48 \cdot \binom{4}{2}} = \frac{52}{3796} = \frac{1}{73}.$$

5. Ottilia gillar att experimentera med trolldeg. Hon blandar i olika mängd av ämnet V till sin mix och testar hur detta påverkar elasticiteten. Hon rullar degen till en boll som hon sedan släpper från ett bord för att se hur högt den studsar tillbaka.

   Hon upprepar experimentet 7 gånger. I tabellen nedan ser vi resultatet av Ottilias experiment.

Mängd V (i gram):	1	2	3	4.5	6	8	10
Studshöjd (i cm):	17.8	21.4	23.6	24	25.8	27	27.3

   Ottilia ansätter en linjär regressionsmodell $y = \beta_0 + \beta_1 x$ där y är studshöjden och x är mängden av ämnet V. Data sammanfattas med att $S_{xx} \approx 64.2$, $S_{yy} \approx 68.3$ och $S_{xy} = 60.6$.

   (a) Skatta β_0 och β_1. (2p)

   (b) Ottilia misstänker att studshöjden ökar precis i takt med mängden V (dvs att $\beta_1 = 1$). Understöder data hennes hypotes? (2p)

   (c) Ange förklaringsgraden och beräkna residualerna. Kommentera ditt resultat. (2p)

Lösning:

(a) Vi har att

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \approx 0.944 \text{ och } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \approx 19.2.$$

(b) Vi betraktar hypteserna $H_0 : \beta_1 = 1$ och $H_1 : \beta_1 \neq 1$. Det enklaste är att skapa ett lämpligt konfidensintervall, och vi väljer (något godtyckligt, andra val är också möjliga) konfidensnivån 95%.

Vi använder att

$$\frac{\hat{\beta}_1 - \beta_1}{s_r / \sqrt{S_{xx}}} \sim t(5)$$

och får med hjälp av tabell att $t_{0.025}(5) \approx 2.571$ så att

$$\begin{aligned} 0.95 &= \mathbb{P}\left(-2.571 \le \frac{\hat{\beta}_1 - \beta_1}{s_r/\sqrt{S_{xx}}} \le 2.571\right) \\ &= \mathbb{P}\left(\hat{\beta}_1 - 2.571\frac{s_r}{\sqrt{S_{xx}}} \le \beta_1 \le \hat{\beta}_1 + 2.571\frac{s_r}{\sqrt{S_{xx}}}\right). \end{aligned}$$

Vi har att

$$s_r^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right) \approx 2.2175$$

så att $s_r = 1.489$. Ett 95% numeriskt K.I. för β_1 blir då

$$I_{\beta_1} = \hat{\beta}_1 \pm 2.571\frac{s_r}{\sqrt{S_{xx}}} \approx [0.47, 1.42].$$

Då $1 \in I_{\beta_1}$ förkastar vi ej H_0 på 95% nivån.

(c) Förklaringsgraden blir

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} \approx 0.8377$$

vilket är ok (not great, not terrible). Residualerna är $e_k = y_k - (\hat{\beta}_0 + \hat{\beta}_1 x_k)$ och vi får att

e_1	e_2	e_3	e_4	e_5	e_6	e_7
-2.33	0.32	1.58	0.56	0.95	0.26	-1.33

Det är mycket anmärkningsvärt att residualerna är negativa på kanterna men alla är positive i mitten. Detta pekar starkt mot att det korrekta sambandet snarare är något i stil med $y = \beta_0 + \beta_1\sqrt{x}$.

6. Låt X vara en slumpvariabel med täthetsfunktion

$$f_X(x) = C_\alpha x^\alpha, \text{ för } 0 \le x \le 1,$$

där $\alpha > 0$ är en parameter.

(a) För varje fixt värde på α, bestäm C_α så att $f_X(x)$ verkligen utgör en täthetsfunktion. (1p)

(b) Låt $Y = X^2 + 1$, bestäm täthetsfunktionen för Y. (3p)

(c) Bestäm täthetsfunktionen för $W = \sqrt{Y}$. (3p)

Lösning:

(a) Vi har att

$$1 = \int_0^1 C_\alpha x^\alpha dx = C_\alpha \left[\frac{x^{1+\alpha}}{1+\alpha}\right]_0^1 = \frac{C_\alpha}{1+\alpha}$$

så att $C_\alpha = 1 + \alpha$.

(b) Vi börjar med fördelningsfunktionen för Y. Vi ser att

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 + 1 \leq y) = \mathbb{P}(X^2 \leq y - 1).$$

Då $X \in [0,1]$ ser vi att $F_Y(y) = 0$ för $y \leq 1$ och $F_Y(y) = 1$ för $y \geq 2$. Vi får också att för $y \in [1,2]$ så är

$$\begin{aligned} F_Y(y) &= \mathbb{P}(X \leq \sqrt{y-1}) \\ &= \int_0^{\sqrt{y-1}} C_\alpha x^\alpha dx = C_\alpha \left[\frac{x^{1+\alpha}}{1+\alpha} \right]_0^{\sqrt{y-1}} = (y-1)^{(1+\alpha)/2}. \end{aligned}$$

Därför blir

$$f_Y(y) = F_Y'(y) = \frac{1+\alpha}{2}(y-1)^{(\alpha-1)/2} \text{ för } y \in [1,2].$$

(c) Man kan antingen använda täthetsfunktionen för Y som vi tagit fram i uppgift (b), eller också uttrycker vi W i termer av X.

Då $Y \in [1,2]$ måste vi ha att $W \in [1,\sqrt{2}]$. Av den anledningen beräknar vi $F_W(w)$ enbart för $w \in [1,\sqrt{2}]$ då $F_W(w)$ måste vara 0 för $w \leq 1$ och 1 för $w \geq \sqrt{2}$. Vi får att för $w \in [1,\sqrt{2}]$

$$\begin{aligned} F_W(w) &= \mathbb{P}(W \leq w) = \mathbb{P}(\sqrt{Y} \leq w) \\ &= \mathbb{P}(Y \leq w^2) = F_Y(w^2) = (w^2-1)^{(1+\alpha)/2}, \end{aligned}$$

så att

$$f_W(w) = F_W'(w) = 2w\frac{1+\alpha}{2}(w^2-1)^{(\alpha-1)/2} = (1+\alpha)w(w^2-1)^{(\alpha-1)/2},$$

för $w \in [1,\sqrt{2}]$.

7. Egil gillar guld och är därför guldgrävare. Han gräver fram guldhaltigt grus som krossas till ett fint damm ur vilket han sedan extraherar själva guldet. Egil använder sig av en traditionell smältmetod, men funderar på att byta till en kemiskt baserad metod (han dränker helt sonika allt damm i cyanid).

   Då den nya metoden är farlig och relativt dyr så vill han veta om den faktiskt ger bättre resultat än den gamla. Han tar därför fram åtta stendammprover från åtta olika markplättar. Varje prov delas sedan upp i två delar och metoderna jämförs.

Prov nr:	1	2	3	4	5	6	7	8
Halt (metod 1, ppm):	6.5	8.1	7.2	4.1	5.9	2.6	7.7	1.6
Halt (metod 2, ppm):	7.4	8.9	8.5	5.4	5.6	3.6	8.5	3.2

   Egil antar att proverna kommer från normalfördelningar.

(a) Sätt upp lämpligt hypotestest för att testa om det är någon skillnad (positiv eller negativ) på signifikansnivån 1 %. (3p)

(b) För att det skall vara lönsamt att byta metod måste skillnaden vara större än 0.3 ppm. Sätt upp ett lämpligt hypotestest för att testa huruvida så är fallet. Använd samma signifikansnivå som i uppgift (a). (3p)

Lösning:

(a) Uppenbarligen rör det sig om ett parat försök (åtta olika prover, varje prov testades med två olika metoder). Vi bildar därför skillnaderna $\Delta_1, \dots, \Delta_8$

Prov nr:	1	2	3	4	5	6	7	8
Skillnad:	0.9	0.8	1.3	1.3	−0.3	1	0.8	1.6

och testar hypotesen $H_0 : \Delta = 0$ mot $H_1 : \Delta \neq 0$.
Vi har att

$$\bar{\Delta} \sim N\left(\Delta, \frac{\sigma^2}{n}\right)$$

där σ^2 är okänd. Vi bildar därför test-statistikan

$$T = \frac{\bar{\Delta}}{s(\Delta)/\sqrt{n}} \sim t(n-1).$$

Vi har enligt tabell att $0.01 = \mathbb{P}(|T| \geq 3.499)$ så att vår förkastningsregion blir $RR = (-\infty, -3.499] \cup [3.499, \infty)$.
Data ger $\bar{\delta} \approx 0.925$ och $s^2 = \frac{1}{7}\sum_{k=1}^{8}(\delta_i - \bar{\delta})^2 \approx 0.325$. Därför blir

$$T(\delta) = \frac{\bar{\delta}}{s/\sqrt{8}} \approx 4.59$$

och vi förkastar därför H_0 på signifikansnivån 1%.

(b) Här vill vi testa $H_0 : \Delta = 0.3$ mot $H_1 : \Delta > 0.3$. Teststatistikan blir (under H_0)

$$T = \frac{\bar{\Delta} - 0.3}{s(\Delta)/\sqrt{n}} \sim t(n-1).$$

Då $0.01 = \mathbb{P}(T \geq 2.998)$ blir vår förkastningsregion RR$= [2.998, \infty)$.
Med insatta data får vi här

$$T(\delta) \approx 3.1$$

och då $T(\delta) \in RR$ så förkastar vi återigen H_0 på signifikansnivån 1%.

8. Låt X ha täthetsfunktion

$$f_X(x) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2} \text{ för } -\infty < x < \infty,$$

där $\alpha > 0$ är okänd.

(a) Hitta momentskattaren (MME:n) för α. (3p)

(b) Hitta maximum likelihood skattaren (MLE:n) för α. (3p)

Lösning:

(a) Vi har att

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \sqrt{\frac{\alpha}{\pi}} x e^{-\alpha x^2} dx = 0$$

av symmetriskäl. Vi måste därför beräkna $\mathbb{E}[X^2]$ och ser då att

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} \sqrt{\frac{\alpha}{\pi}} x^2 e^{-\alpha x^2} dx \\ &= \sqrt{\frac{\alpha}{\pi}} \left[x \frac{e^{-\alpha x^2}}{-2\alpha} \right]_{-\infty}^{\infty} + \frac{1}{2\alpha} \int_{-\infty}^{\infty} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2} dx = 0 + \frac{1}{2\alpha}, \end{aligned}$$

där vi utnyttjar att integranden är en täthetsfunktion (så att integralen blir 1). Momentmetoden säger då att vi skall sätta

$$\overline{X^2} = \frac{1}{2\hat{\alpha}} \quad \Rightarrow \quad \hat{\alpha} = \frac{1}{2\overline{X^2}}.$$

(b) Vi bildar likelihooden

$$Lik(\alpha) = \prod_{k=1}^{n} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha X_k^2} = \left(\frac{\alpha}{\pi}\right)^{n/2} e^{-\alpha \sum_{k=1}^{n} X_k^2}$$

och sedan log-likelihooden

$$l(\alpha) = \frac{n}{2}(\log \alpha - \log \pi) - \alpha \sum_{k=1}^{n} X_k^2,$$

så att

$$l'(\alpha) = \frac{n}{2\alpha} - \sum_{k=1}^{n} X_k^2 = 0 \quad \Rightarrow \quad \hat{\alpha} = \frac{n}{2\sum_{k=1}^{n} X_k^2} = \frac{1}{2\overline{X^2}}.$$

Dessutom har vi att $l''(\alpha) = \frac{-n}{2\alpha^2} < 0$ så vi har hittat vårt maximum.