STOCKHOLMS UNIVERSITET
Statistiska institutionen
Jessica Franzén

# TENTAMEN I STATISTISK TEORI MED TILLÄMPNINGAR II
2018-11-26

---

**Skrivtid:** 12.00-17.00
**Godkända hjälpmedel:** Miniräknare, språklexikon.
Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.

---

OBS! Glöm inte att ange nödvändiga antaganden där det behövs.

**Uppgift 1.** (20 poäng)
Nedan följer ett antal påståenden, avgör om dessa är sanna eller falska eller om mer information behövs. Motivera ditt svar.

i) Enligt centrala gränsvärdessatsen närmar sig fördelningen för populationen en normalfördelning när urvalsstorleken ökar.

ii) Enligt centrala gränsvärdessatsen närmar sig fördelningen för urvalsmedelvärdet en standardiserad normalfördelning när urvalsstorleken ökar.

iii) Väntevärdet i samplingfördelningen för urvalsmedelvärdet närmar sig populationsmedelvärdet när urvalsstorleken ökar.

iv) Urvalsmedelvärdet är en konsistent estimator för populationsmedelvärdet.

v) Estimator A är mer effektiv än estimator B om estimator A:s varians är lägre än estimator B:s varians.

**Uppgift 2.** (20 poäng)
Man vill testa effekten av ett program vars syfte är att öka den sociala kompetensen hos de deltagande personerna. 15 personer gick programmet och man mätte den sociala kompetensen för varje person på en sjugradig skala (högre värde motsvarar högre social kompetens) före och efter programmet.

| Person | Social kompetens före programmet | Social kompetens efter programmet |
|---|---|---|
| 1 | 5 | 6 |
| 2 | 3 | 2 |
| 3 | 4 | 4 |
| 4 | 2 | 4 |
| 5 | 1 | 3 |
| 6 | 6 | 6 |
| 7 | 7 | 7 |
| 8 | 3 | 5 |
| 9 | 2 | 1 |
| 10 | 3 | 5 |
| 11 | 5 | 5 |
| 12 | 1 | 3 |
| 13 | 4 | 4 |
| 14 | 4 | 5 |
| 15 | 3 | 2 |

a) Testa med ett teckentest om programmet har positiv effekt d.v.s. om den sociala kompetensen ökar. Räkna ut p-värdet.

b) Bestäm testets signifikansnivå om man bestämmer sig för att förkasta nollhypotesen om minst 8 av 10 personer får ett högre testresultat efter programmet jämfört med före programmet.

c) Bestäm sannolikheten för ett typ II-fel om $p = 0.8$ och samma förutsättningar som i b) gäller.

**Uppgift 3.** (20 poäng)
Hur lång tid en tennismatch pågår varierar mycket. Nedan ges matchlängden i minuter för 16 slumpmässigt valda matcher i en Wimbledonturnering.

73, 76, 59, 104, 114, 106, 79, 74, 142, 129, 95, 56, 84, 142, 106, 75

a) Konstruera ett 95%-igt konfidensintervall för standardavvikelsen av matchlängderna i Wimbledonturneringen.

b) Konstruera två enkelsidiga 95%-iga konfidensintervall för standardavvikelsen av matchlängderna i Wimbledonturneringen.

**Uppgift 4.** (20 poäng)
Låt $Y_1$, $Y_2$, $Y_3$ vara ett slumpmässigt urval från

$$f(y) = \begin{cases} \frac{2y}{\theta^2}, & 0 \leq y \leq \theta \\ 0, & \text{annars.} \end{cases}$$

$\widehat{\theta} = \frac{Y_1+Y_2+Y_3}{2}$ är en estimator för $\theta$.

a) Bestäm väntevärdet för $\widehat{\theta}$. Är $\widehat{\theta}$ väntevärdesriktig?

b) Bestäm variansen för $\widehat{\theta}$.

**Uppgift 5.** (20 poäng)
I tillverkningsprocessen av en viss produkt produceras dessvärre en andel defekta produkter. Antalet producerade produkter $Y$ tills man får en defekt produkt antas följa en geometriskt fördelning, där parametern $p$ är sannolikheten för en defekt produkt.

a) Bestäm maximum likelihood-skattningen av parametern $p$.

b) Bestäm teststatistikan för likelihoodkvottestet för att testa $H_0 : p = 0.3$ mot $H_A : p \neq 0.3$ då man gör $n$ stycken observationer och för varje observation noterar $Y$.

c) Ange kritiskt område för testet i b) och genomför testet med följande data: man har gjort $n = 35$ observationer och medelvärdet för dessa 35 observationer är $\overline{y} = 3.75$. Använd signifikansnivån 0.05.

Statistiska institutionen

Stockholms universitet

# Rättningsblad

**Datum:** 26/11/18

**Sal:** Värtasalen

**Tenta:** Statistisk teori med tillämpningar
**Kurs:** Statistisk teori med tillämpningar II

**ANONYMKOD:** 0029-XOF

☒ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

**OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN**

**Markera besvarade uppgifter med kryss**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Antal inl. blad |
|---|---|---|---|---|---|---|---|---|---|
| X | X | X | X | X | | | | | 6 |
| Lär.ant. 13 | 17 | 10 | 20 | 10 | | | | | |

| POÄNG | BETYG | Lärarens sign. |
|---|---|---|
| 70+0=70 | C | |

SU, STATISTIK

Skrivsal: Värta Anonymkod: 0029-XOF Blad nr: 1

1) a) Falskt. Populationen har alltid dess ursprungliga fördelning, oavsett urvalsstorlek den lofs närmar sig en normalfördelning i det fall den ursprungliga också är en normalfördelning

4

b) Falskt. Fördelningen närmar sig en approximativ normalfördelning, dock ej nödvandigtvis en standardiserad

3

iii) Sant. Eftersom variansen minskar när n ökar

0

iv) Sant. Eftersom både variansen och bias går mot 0 när n ökar.

2

behövs mer info

v) Sant, men det förutsätter också att både estimator A och estimator B är Väntvärdesriktiga.

4

**SU, STATISTIK**

Skrivsal: Värta Anonymkod: 0029-XOF Blad nr: 2

2)

| Person | X | Y | $D_i$ (X-Y) | Tecken |
|---|---|---|---|---|
| 1 | 5 | 6 | -1 | - |
| 2 | 3 | 2 | 1 | + |
| 3 | 4 | 4 | 0 | Tie |
| 4 | 2 | 4 | -2 | - |
| 5 | 1 | 3 | -2 | - |
| 6 | 6 | 6 | 0 | Tie |
| 7 | 7 | 7 | 0 | Tie |
| 8 | 3 | 5 | -2 | - |
| 9 | 2 | 1 | 1 | + |
| 10 | 3 | 5 | -2 | - |
| 11 | 5 | 5 | 0 | Tie |
| 12 | 1 | 3 | -2 | - |
| 13 | 4 | 4 | 0 | Tie |
| 14 | 4 | 5 | -1 | - |
| 15 | 3 | 2 | 1 | + |

X = social kompetens före programmet

Y = Social kompetens efter

Parvisa observationer, vi antar oberoende mellan paren och s.u.

$H_0: p = 0,5$

$H_a: p > 0,5$ K

Testvariabel:

$M \sim Bin(10, 0.5)$ om Ho är sann

m = antal ~~negativa~~ pos differenser = 7

$P(M \geq 7) = 1 - P(M \leq 6) = 1 - 0,82813 = 0,17187$ = p-värde

9

Svar: Med ett p-värde på 0,17187 kan vi inte på nån rimlig signifikansnivå dra slutsatsen att kompetensen ökar genom programmet

b) $\alpha = P(M \geq 8) = 1 - P(M \leq 7) = 1 - 0,94531 = 0,0547$

Svar: testets sign.n. är 0,0547.

5

**SU, STATISTIK**

**Skrivsal:** Värta **Anonymkod:** 0029-XOF **Blad nr:** 3

$\beta$

c) Typ II-fel: Acceptera $H_0 \mid H_A$ är sann ($H_0$ är falsk)

Det innebär att ~~minst~~ max 7 av 10 måste få högre provresultat vid $p = 0{,}8$.

Pga likformighet kan vi uttrycka det som sannolikheten $1 - P(M \leq 2)$ vid $p = 0{,}2$ och $n = 10$

Läst ur tabell blir det p-värde $= 0{,}6778$, och därmed också $\beta$

$1 - 0{,}6778 = 0{,}3222$

Svar: $\beta = 0{,}678$ = sannolikheten för typ II-fel

3

**SU, STATISTIK**

**Skrivsal:** Värta **Anonymkod:** 0029-XOF **Blad nr:** 4

3) $\bar{Y} = (73+76+59+104+114+106+79+74+142+129+95+56+84+142+106+75)/16 = 94,625$ min

Vi antar att matchlängden är normalfördelad och att observationerna är oberoende och med samma väntvärde och varians.

$$S^2 = \frac{\sum_{i=1}^{n} Y_i^2 - n(\bar{Y})^2}{n-1} = \frac{(73^2+76^2+59^2+104^2+114^2+106^2+79^2+74^2+142^2+129^2+95^2+56^2+84^2+142^2+106^2+75^2) - 16\cdot(94,625)^2}{16-1}$$

$S^2 = 742,4$

K.i. för variansen är enligt formelblad

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} ; \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}} \quad \frac{(16-1)\,742,4}{27,49} ; \frac{(16-1)\,742,4}{6,26}$$

$= [405,1 ; 1778,9]$

för att erhålla k.i för standardavvikelsen tar vi roten ur för de k.i värden vi fick för variansen

$\sqrt{405,1}$, $\sqrt{1778,9}$ $\Rightarrow$ $[20,13 ; 42,18]$ = Svar = k.i för matchlängdernas

10

**SU, STATISTIK**

Skrivsal: Värta Anonymkod: 0029-XOF Blad nr: 5

4) a) $\hat{\theta} = \dfrac{Y_1 + Y_2 + Y_3}{2} \overset{\text{pga oberoende}}{=} \dfrac{E(Y_1) + E(Y_2) + E(Y_3)}{2}$

$f(y) = \begin{cases} \dfrac{2y}{\theta^2}, & 0 \le y \le \theta \\ 0, & \text{annars} \end{cases}$

$E(Y) = \int y \cdot f(y)\, dy = \int_0^\theta y \cdot \dfrac{2y}{\theta^2} = \int_0^\theta \dfrac{2y^2}{\theta^2} = \left[\dfrac{2y^3}{3\theta^2}\right]_0^\theta =$

$\dfrac{2\theta^3}{3\theta^2} - 0 = \dfrac{2\theta}{3}$ sätter in i $\hat{\theta}$ ger att

$\hat{\theta} = \dfrac{\frac{2\theta}{3} + \frac{2\theta}{3} + \frac{2\theta}{3}}{2} = \dfrac{2\theta}{2} = \theta$

10

Svar: Ja, $\hat{\theta}$ är väntvärdesriktig

b) $V\hat{\theta} \overset{\text{pga oberoende}}{=} \dfrac{1}{4}\left(V(Y_1) + V(Y_2) + V(Y_3)\right)$

$V(Y) = \int y^2 f(y)\, dy - [E(Y)]^2 = \Big/ \int y^2 \dfrac{2y}{\theta^2} = \int \dfrac{2y^3}{\theta^2} = \left[\dfrac{2y^4}{4\theta^2}\right]_0^\theta =$

$\dfrac{2\theta^2}{4} - 0 = \dfrac{\theta^2}{2} \Big/ = \dfrac{\theta^2}{2} - \left(\dfrac{2\theta}{3}\right)^2 = \dfrac{\theta^2}{2} - \dfrac{4\theta^2}{9} = \dfrac{\theta^2}{18}$

sätter in i $V\hat{\theta}$ ger att

$V(\hat{\theta}) = \dfrac{1}{4}\left(\dfrac{\theta^2}{18} + \dfrac{\theta^2}{18} + \dfrac{\theta^2}{18}\right) = \dfrac{\theta^2}{6}\left(\dfrac{1}{4}\right) = \dfrac{\theta^2}{24}$

10

Svar: $V(\hat{\theta}) = \dfrac{\theta^2}{24}$

**SU, STATISTIK**

Skrivsal: Värta  Anonymkod: 0029-XOF  Blad nr: 6

5) c)

$$\Lambda_{LR} = \frac{0{,}3^{35}(1-0{,}3)^{\overset{n\cdot\bar{y}=\sum y_i}{35\cdot 3{,}75-35}}}{\left(\frac{1}{3{,}75}\right)^{35}\left(1-\frac{1}{3{,}75}\right)^{35\cdot 3{,}75-35}} = 0{,}70$$ K

$$W = -2\ln(\Lambda_{LR}) \sim \chi^2(1) = 0{,}71$$ K

Kritisk värde fås ur tabell för $\chi^2$-förd med $\alpha = 0{,}05$ och 1 frihetsgrad. $= 3{,}84 = k$

Vårt observerade värde $W = 0{,}71$ är mindre än k viket gör att vi ej! kan förkasta $H_0$. K

(10)

Svar: Vi kan ej förkasta $H_0$

Var är fråga 5 a) och b)?

Jag misstänker att du missat att lägga med dem.