

S0008M-2024-08-27

19 okt. 2026, 15 min läsning

UTKAST

#tenta #sannolikhetslära #statistik

Kurs: S0008M — Sannolikhetslära och statistik **Datum:** 2024-08-27 **Källa:** originaltentamen ·
lärarens lösningar (bara uppgift 1–13)

Översikt

Del 1 (uppgift 1–12, bara svar):

- betingning med oberoende händelser (1), standardavvikelse för en nettovinst (2) och hypergeometrisk fördelning (3)
- α ur en given varians (4), en dubbelsidig exponentialtätthet (5) och “inom en standardavvikelse” för $R(0, 2)$ (6)
- CGS för antal bilar (7), styrka för ett binomialtest (8) och frihetsgrader för två stickprov (9)
- teckentest (10), mötesproblemet (11) och en differens i tvådimensionell normalfördelning (12)

Del 2 (uppgift 13–15):

- väntevärdet av antal jobb utan oberoende (13)
- minsta värdet som skattning i $R(\theta, 0)$ (14)
- sannolikheten att två konfidensintervall överlappar (15)

Så använder du sidan

Försök själv först. Ledtråd 1 säger vilken sorts uppgift det är, ledtråd 2 (inuti ledtråd 1) ger modellen och ledtråd 3 (inuti ledtråd 2) nyckelsteget. **Lösning** är hela uträkningen. Lärarens facit täcker uppgift 1–13; uppgift 14 och 15 är lösta två gånger oberoende av varandra.

1. Betingat på $A \cup B^c$

För de **oberoende** händelserna A och B gäller att $P(A) = 0,3$, $P(B) = 0,74$. Låt $D = A \cup B^c$. Beräkna den betingade sannolikheten $P(A \mid D)$. Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Betingad sannolikhet: $P(A \mid D) = P(A \cap D)/P(D)$.

Ledtråd 2

$A \subset D$, så $A \cap D = A$ (Utfallsrum och händelser).

Ledtråd 3

$P(D) = P(A) + P(B^c) - P(A)P(B^c)$, eftersom A och B^c också är **oberoende**.

Lösning

Eftersom $A \subset D$ är $P(A \cap D) = P(A) = 0,3$. Additionssatsen och oberoendet ger

$$P(D) = P(A) + P(B^c) - P(A)P(B^c) = 0,3 + 0,26 - 0,3 \cdot 0,26 = 0,482$$

$$P(A \mid D) = \frac{0,3}{0,482} \approx \boxed{62,24 \%}$$

Kontrollerat mot facit.

2. Karls nettovinst

Karl jobbar som inkastare på en bar på en grekisk ö, vilket innebär att han försöker få förbipasserade turister att besöka baren. För varje person han lyckas få att gå in i baren får han 35 kr. Samtidigt har Karl kostnader på 300 kr per dag. Så om han till exempel lyckas locka 20 kunder en dag blir alltså hans

nettovinst den dagen lika med 400.

Antag att antalet kunder som Karl lyckas locka till baren på en dag har en Poissonfördelning och att han i genomsnitt lyckas locka 50 kunder på en dag. Vad är då **standardavvikelsen** för hans nettovinst på en dag? Svara med två decimaler.

Totalpoäng: 2

Ledtråd 1

Nettovinsten är en **linjär funktion** av antalet kunder.

Ledtråd 2

$\eta = 35\xi - 300$ med $\xi \in \text{Po}(50)$. För en **Poissonfördelning** är $V(\xi) = \lambda$.

Ledtråd 3

Konstanten -300 påverkar inte spridningen: $D(\eta) = 35 D(\xi)$.

Lösning

Med $\xi \in \text{Po}(50)$ är $V(\xi) = 50$. Nettovinsten $\eta = 35\xi - 300$ har

$$V(\eta) = 35^2 \cdot 50, \quad D(\eta) = 35\sqrt{50} \approx \boxed{247,49 \text{ kr}}$$

Kontrollerat mot facit.

3. Högst ett defekt kort i urvalet

Vid etsning av kretskort är andelen defekta ofta hög, och därför kontrolleras de färdiga korten. Kort läggs ihop i förpackningar om 20 kort. 6 av dessa ska tas ut för undersökning. Om det bland de 20 finns 5 defekta, hur stor är sannolikheten att det i urvalet finns högst 1 defekt kort? Svara i procent med två decimaler.

Totalpoäng: 2

 Ledtråd 1

Dragning utan återläggning ur en ändlig mängd: **Hypergeometrisk fördelning**.

 Ledtråd 2

ξ = antal defekta i urvalet, $\xi \in \text{Hyp}(20, 6, 0,25)$.

 Ledtråd 3

$$P(\xi \leq 1) = P(\xi = 0) + P(\xi = 1), \text{ med } P(\xi = k) = \frac{\binom{5}{k} \binom{15}{6-k}}{\binom{20}{6}}.$$

 Lösning

$$P(\xi \leq 1) = \frac{\binom{5}{0} \binom{15}{6} + \binom{5}{1} \binom{15}{5}}{\binom{20}{6}} = \frac{5005 + 5 \cdot 3003}{38760} = \frac{20020}{38760} \approx \boxed{51,65 \%}$$

Kontrollerat mot facit.

4. För vilket a blir standardavvikelsen 4?

Betrakta följande sannolikhetsfördelning där konstanten a är strikt positiv.

x	$-a$	0	a
$P(\xi = x)$	0,25	0,5	0,25

För vilket värde på a blir standardavvikelsen i fördelningen lika med 4? Det vill säga, för vilket värde på a gäller att $V(\xi) = 16$? Svara med två decimaler.

Totalpoäng: 2

 Ledtråd 1

Räkna ut **variansen** direkt ur definitionen.

 Ledtråd 2

Fördelningen är symmetrisk kring 0, så $E(\xi) = 0$ och $V(\xi) = E(\xi^2)$.

 Ledtråd 3

$$E(\xi^2) = a^2 \cdot 0,25 + a^2 \cdot 0,25 = 0,5 a^2.$$

 Lösning

Symmetrin ger $E(\xi) = 0$, så

$$V(\xi) = (-a)^2 \cdot 0,25 + 0^2 \cdot 0,5 + a^2 \cdot 0,25 = 0,5 a^2$$

$0,5 a^2 = 16$ ger $a^2 = 32$ och

$$a = \sqrt{32} \approx \boxed{5,66}$$

Kontrollerat mot facit.

5. Dubbelsidig exponentialtätet

Slumpvariabeln ξ har en kontinuerlig fördelning vars frekvensfunktion är $f(x) = 0,5 \cdot e^{-|x|}$, $-\infty < x < \infty$. Beräkna sannolikheten $P(-1 < \xi < 1)$. Svara i procent med en decimal.

Totalpoäng: 2

 Ledtråd 1

Sannolikhet ur en **frekvensfunktion**: integrera f över intervallet.

 Ledtråd 2

f är symmetrisk kring 0, så $P(-1 < \xi < 1) = 2P(0 < \xi < 1)$.

 Ledtråd 3

För $x > 0$ är $|x| = x$: integrera $0,5 e^{-x}$ från 0 till 1.

 Lösning

$$P(0 < \xi < 1) = \int_0^1 0,5 e^{-x} dx = 0,5(1 - e^{-1})$$

Symmetrin ger

$$P(-1 < \xi < 1) = 2 \cdot 0,5(1 - e^{-1}) = 1 - e^{-1} \approx \boxed{63,2 \%}$$

Kontrollerat mot facit.

6. Inom en standardavvikelse i $R(0, 2)$

I normalfördelningen återfinns ca 68 % av massan inom en standardavvikelse från väntevärdet. Mer exakt gäller att om $\xi \in N(\mu, \sigma)$ så har vi $P(\mu - \sigma \leq \xi \leq \mu + \sigma) = 0,6826$.

Beräkna motsvarande sannolikhet för $R(0, 2)$ -fördelningen. Det vill säga, beräkna $P(\mu - \sigma \leq \xi \leq \mu + \sigma)$ om $\xi \in R(0, 2)$. Svara i procent med två decimaler.

Totalpoäng: 2

 Ledtråd 1

Likformig fördelning: sannolikheten för ett delintervall är dess längd delad med hela längden.

 Ledtråd 2

$R(a, b)$ har $\mu = (a + b)/2$ och $\sigma = (b - a)/\sqrt{12}$.

 Ledtråd 3

Här är $\mu = 1$ och $\sigma = 2/\sqrt{12}$. Intervallet $[\mu - \sigma, \mu + \sigma]$ har längden 2σ och ligger inom $[0, 2]$.

 Lösning

$\mu = 1$ och $\sigma = 2/\sqrt{12} \approx 0,577$. Intervallet $[\mu - \sigma, \mu + \sigma]$ ligger inom $[0, 2]$, så

$$P(\mu - \sigma \leq \xi \leq \mu + \sigma) = \frac{2\sigma}{2} = \frac{2}{\sqrt{12}} = \frac{1}{\sqrt{3}} \approx \boxed{57,74 \%}$$

Kontrollerat mot facit. Lärarens facit anger 0,577, alltså samma värde med färre decimaler.

7. Högst 415 bilar i bostadsområdet

Ett bostadsområde planeras för 500 hushåll. En undersökning visar att antalet bilar per hushåll är 0, 1 eller 2 med sannolikheten 0,3, 0,6 respektive 0,1. Vad är sannolikheten att det totalt blir högst 415 bilar i bostadsområdet? Svara i procent med minst en decimal.

Kommentar: Som vanligt får approximationer användas om sådana skulle kunna motiveras.

Totalpoäng: 2

 Ledtråd 1

En summa av 500 oberoende hushåll: **Centrala gränsvärdessatsen**.

 Ledtråd 2

Räkna ut $[V\text{äntevärde}|E(\xi_i)]$ och $V(\xi_i)$ för ett hushåll ur tabellen.

 Ledtråd 3

$E(\xi_i) = 0,8$ och $V(\xi_i) = 1,0 - 0,8^2 = 0,36$. Summan är $\approx N(400, \sqrt{500} \cdot 0,6)$.

 Lösning

För ett hushåll:

$$E(\xi_i) = 0 \cdot 0,3 + 1 \cdot 0,6 + 2 \cdot 0,1 = 0,8, \quad E(\xi_i^2) = 0,6 + 4 \cdot 0,1 = 1,0$$

$$V(\xi_i) = 1,0 - 0,8^2 = 0,36, \quad D(\xi_i) = 0,6$$

Enligt CGS är $\sum_{i=1}^{500} \xi_i \approx N(400, \sqrt{500} \cdot 0,6) = N(400; 13,42)$, så

$$P\left(\sum_{i=1}^{500} \xi_i \leq 415\right) \approx \Phi\left(\frac{415 - 400}{13,42}\right) = \Phi(1,118) \approx \boxed{86,8 \%}$$

På räknaren. `normCdf(-1E99, 415, 400, sqrt(500)*0.6)`.

Kontrollerat mot facit. Lärarens facit skriver $\mu = 0,6$ och summerar till 100 i sista raden, men det är skrivfel: räkningen använder $N(400, \sqrt{500} \cdot 0,6)$, alltså $\mu = 0,8$ och 500 hushåll. Med halvkorrektion (415,5) blir svaret 87,6 %.

8. Styrkan för ett test av blixtar

Ljusblixtar som levereras till ett laboratorium uppges ha sannolikheten $p = 0,03$ att inte fungera. För att testa $H_0 : p = 0,03$ mot $H_1 : p > 0,03$ provar man 60 blixtar. Om antalet felaktiga är 6 eller flera förkastas H_0 .

Beräkna testets styrka då $p = 0,075$. Svara i procent med en decimal.

Totalpoäng: 2

 Ledtråd 1

Styrkan är sannolikheten att förkasta H_0 när ett visst alternativ är sant.

 Ledtråd 2

ξ = antal felaktiga av 60, $\xi \in \text{Bin}(60; p)$ (**Binomialfördelning**).

 Ledtråd 3

Styrkan är $P(\xi \geq 6 \mid p = 0,075) = 1 - P(\xi \leq 5 \mid p = 0,075)$.

 Lösning

Med $\xi \in \text{Bin}(60; 0,075)$:

$$S(0,075) = P(\xi \geq 6) = 1 - \sum_{k=0}^5 \binom{60}{k} 0,075^k 0,925^{60-k} \approx \boxed{29,4 \%}$$

Poissonapproximationen $\text{Po}(4,5)$ ger 29,7 %.

På räknaren. `1-binomCdf(60,0.075,5)`.

Kontrollerat mot facit.

9. Frihetsgrader för två stickprov

Följande värden är ett observerat stickprov från en $N(\mu_1, \sigma_1)$ -fördelning (μ_1, σ_1 är okända):

6,4	6,1	5,6	5	6,8	7	5,1
-----	-----	-----	---	-----	---	-----

Följande värden är ett observerat stickprov från en $N(\mu_2, \sigma_2)$ -fördelning (μ_2, σ_2 är okända):

5,2	3,5	5,6	5,7	5	3,6	6
-----	-----	-----	-----	---	-----	---

De två stickproven är oberoende och $\sigma_1 = \sigma_2$. För att beräkna ett 90 % konfidensintervall för $\mu_1 - \mu_2$ behöver man använda en lämplig t -fördelning. Hur många frihetsgrader har den t -fördelningen? Svara med ett heltal.

Totalpoäng: 2

Ledtråd 1

Två stickprov med gemensam okänd σ .

Ledtråd 2

Den sammanvägda skattningen σ^* bygger på båda stickprovens kvadratsummor.

Ledtråd 3

Varje stickprov förlorar en frihetsgrad på sitt eget medelvärde: $n_1 + n_2 - 2$.

Lösning

Med gemensam okänd σ används t -fördelningen med

$$f = n_1 + n_2 - 2 = 7 + 7 - 2 = \boxed{12}$$

frihetsgrader. Mätvärdena behövs inte för att svara.

Kontrollerat mot facit.

10. Teckentest för reaktionstid

En forskare vill undersöka om en viss medicin har en negativ effekt på reaktionstiden vid bilkörning. För att undersöka om så är fallet valde forskaren ut 10 bilförare slumpmässigt och mätte deras reaktionstider (enhet: millisekunder) före och efter att de tagit medicinen. Resultatet visas i tabellen nedan.

Reaktionstid före	273,7	259,3	259,4	231,1	227	209,8
Reaktionstid efter	282,8	256,4	258,1	239,8	230,6	214,5

Forskaren beslutade sig för att man i denna situation *inte* kunde anta att reaktionstiderna och deras differenser var normalfördelade. För att testa

H_0 : medicinen påverkar inte reaktionstiden i genomsnitt

mot den tvåsidiga mothypotesen

H_1 : reaktionstiden är i genomsnitt längre eller kortare då medicinen används

valde forskaren därför att använda metoden med teckentest. Vilken är den lägsta signifikansnivå på vilken H_0 kan förkastas, det vill säga vad blir P-värdet α_0 ? Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Stickprov i par utan normalfördelning: teckenmetoden (samma idé som i **Teckenintervall**) med **direktmetoden**.

Ledtråd 2

Räkna differenserna efter – före. Under H_0 är varje tecken positivt eller negativt med sannolikheten 0,5: $\xi \in \text{Bin}(10; 0,5)$.

Ledtråd 3

Tre differenser är negativa. Tvåsidig mothypotes: $\alpha_0 = 2P(\xi \leq 3)$.

Lösning

Differenserna efter – före är

9,1; -2,9; -1,3; 8,7; 3,6; 4,5; 8,2; 9,9; 5,6; -2,4

Låt ξ = antal förare med kortare reaktionstid efteråt. Under H_0 är $\xi \in \text{Bin}(10; 0,5)$, och vi har observerat $x = 3$. Mothypotesen är tvåsidig, så

$$\alpha_0 = 2P(\xi \leq 3) = 2 \cdot \frac{1 + 10 + 45 + 120}{1024} = \frac{352}{1024} \approx \boxed{34,38 \%}$$

H_0 kan alltså inte förkastas på någon vanlig signifikansnivå.

På räknaren. `2*binomCdf(10,0.5,3)`.

Kontrollerat mot facit. Lärarens facit skriver också " $= 1 - P(\xi \leq 1)$ " i mellanledet, vilket är ett skrivfel; slutvärdet 0,3438 stämmer.

11. Mötesproblemet

Två personer stämmer möte på ett kafé "någon gång mellan kl. 13:00 och kl. 14:00". Antag att personernas ankomsttider kan beskrivas med en rektangelfördelning, och att de väntar 40 minuter och sedan går. Vad är då sannolikheten att de träffas? Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Två oberoende **likformiga** ankomsttider ger en likformig fördelning på en kvadrat (Flerdimensionella fördelningar).

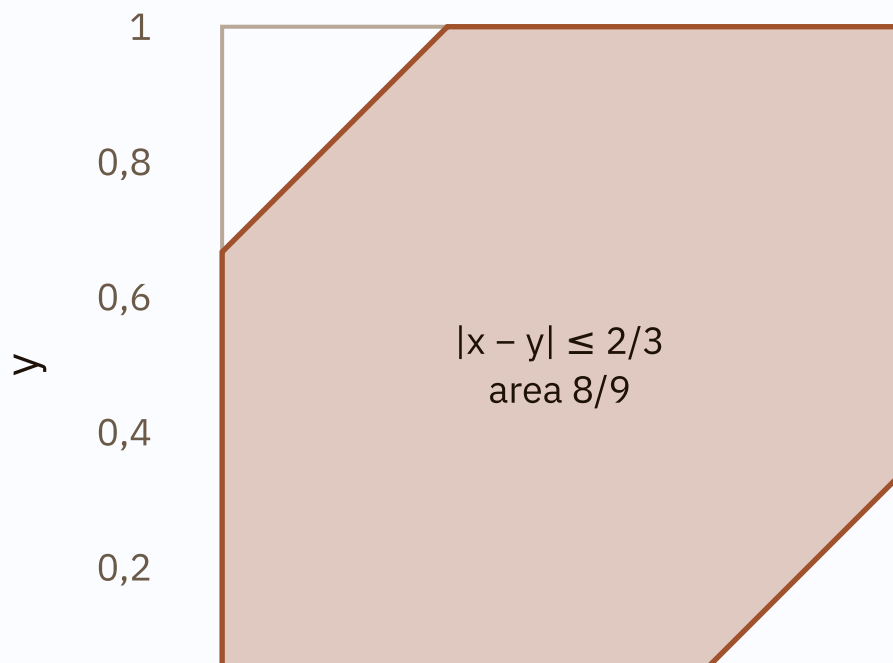
Ledtråd 2

Mät tiden i timmar efter 13:00: $\xi, \eta \in R(0, 1)$. De träffas om $|\xi - \eta| \leq 2/3$.

Ledtråd 3

Sannolikheten är arean av bandet kring diagonalen. Det är enklare att dra bort de två hörntrianglarna.

 Lösning



Med tiden i timmar efter 13:00 är (ξ, η) likformigt fördelad på enhetskvadraten, och de träffas om $|\xi - \eta| \leq 2/3$. Utanför bandet ligger två trianglar med kateterna $1/3$:

$$P(|\xi - \eta| \leq 2/3) = 1 - 2 \cdot \frac{1}{2} \left(\frac{1}{3}\right)^2 = 1 - \frac{1}{9} = \frac{8}{9} \approx \boxed{88,89 \%}$$

Kontrollerat mot facit.

12. Överstiger Majas utgifter Joels?

Majas och Joels utgifter för kursmaterial (enhet: kr) kan betraktas som slumpmässiga, där Majas utgifter per månad har väntevärdet 480 kr och standardavvikelsen 31 kr, medan Joels utgifter per månad antas ha väntevärdet 465 kr och standardavvikelsen 24 kr. Antag att Majas och Joels utgifter för kursmaterial är beroende, med korrelation lika med 0,95, samt att de kan beskrivas med en tvådimensionell normalfördelning.

Vad är sannolikheten att Majas utgifter under en månad överstiger Joels? Ange ditt svar i procent med två decimaler.

Totalpoäng: 3

Ledtråd 1

Studera differensen $\xi = \xi_M - \xi_J$. I en tvådimensionell normalfördelning är den normalfördelad.

Ledtråd 2

$$V(\xi_M - \xi_J) = \sigma_M^2 + \sigma_J^2 - 2\rho\sigma_M\sigma_J \text{ (Kovarians).}$$

Ledtråd 3

$$E(\xi) = 15 \text{ och } V(\xi) = 961 + 576 - 1413,6. \text{ Sök } P(\xi > 0).$$

Lösning

$$E(\xi) = 480 - 465 = 15, \quad V(\xi) = 31^2 + 24^2 - 2 \cdot 0,95 \cdot 31 \cdot 24 = 123,4$$

så $\xi \in N(15; 11,11)$ och

$$P(\xi > 0) = 1 - \Phi\left(\frac{0 - 15}{11,11}\right) = \Phi(1,350) \approx \boxed{91,15 \%}$$

Den höga korrelationen gör differensen liten i spridning, trots att var och en har stor spridning.

På räknaren. `normCdf(0,1E99,15,sqrt(123.4))`.

Kontrollerat mot facit. Lärarens facit skriver $\sqrt{1234}$, men det är ett skrivfel för $\sqrt{123,4}$; svaret 0,9115 bygger på det rätta värdet.

13. Del 2: väntevärdet av antal jobb

En entreprenör har lämnat in anbud på tre olika jobb a , b och c . Hon bedömer sannolikheten att få de olika jobben till 0,3 för a , 0,6 för b och 0,75 för c . Händelserna att få olika jobb kan *inte* antas oberoende. Låt ξ vara det totala antalet jobb som entreprenören får (av dessa tre).

Beräkna väntevärdet av ξ .

Totalpoäng: 10

Ledtråd 1

Fördelningen för ξ går inte att bestämma utan oberoende. Men **väntevärdet** av en summa kräver inget oberoende.

Ledtråd 2

Skriv $\xi = \xi_a + \xi_b + \xi_c$, där $\xi_a = 1$ om hon får jobb a och 0 annars.

Ledtråd 3

$$E(\xi_a) = 0 \cdot P(\text{inte } a) + 1 \cdot P(a) = P(a).$$

Lösning

Modell. Låt $\xi_a = 1$ om entreprenören får jobb a och $\xi_a = 0$ annars, och definiera ξ_b , ξ_c på samma sätt. Då är

$$\xi = \xi_a + \xi_b + \xi_c$$

Väntevärden. För en variabel som bara tar värdena 0 och 1 är väntevärdet sannolikheten för värdet 1:

$$E(\xi_a) = 0,3, \quad E(\xi_b) = 0,6, \quad E(\xi_c) = 0,75$$

Summan. Väntevärdet av en summa är summan av väntevärdena, oavsett beroende (Linjära funktioner av stokastiska variabler):

$$E(\xi) = 0,3 + 0,6 + 0,75 = \boxed{1,65}$$

Beroendet påverkar variansen och fördelningen för ξ , men inte väntevärdet.

Kontrollerat mot facit.

14. Del 2: minsta värdet som skattning i $R(\theta, 0)$

Antag att $x_1, \dots, x_n$ är ett stickprov från $R(\theta, 0)$ där $\theta < 0$ är okänd.

- a) Förklara varför det minsta värdet, det vill säga $\min(x_1, \dots, x_n)$, är en rimlig skattning av θ .
- b) Förklara varför det minsta värdet, det vill säga $\min(x_1, \dots, x_n)$, inte blir en väntevärdesriktig skattning av θ .

På b) får du gärna bevisa påståendet matematiskt, men det är inte nödvändigt.

Totalpoäng: 10

Ledtråd 1

Punktskattning: θ är fördelningens vänstra ändpunkt. Var hamnar observationerna i förhållande till θ ?

Ledtråd 2

Varje $x_i \geq \theta$. Väntevärdesriktig betyder $E(\theta^*) = \theta$.

Ledtråd 3

För $\theta < y < 0$ är $P(\xi_i > y) = y/\theta$, så $P(\min > y) = (y/\theta)^n$. Använd $E(\eta) = \theta + \int_{\theta}^0 P(\eta > y) dy$ för en variabel η på $[\theta, 0]$.

Lösning

a) Rimlig skattning. [[Likformig fördelning| $R(\theta, 0)$]] har sin massa på $[\theta, 0]$, så alla observationer ligger till höger om θ . Det minsta värdet är den observation som ligger närmast θ . Med många observationer är det mycket troligt att någon hamnar nära den vänstra ändpunkten, så $\min$ närmar sig θ när n växer. Dessutom kan θ inte vara större än $\min(x_1, \dots, x_n)$, eftersom inga värden under θ är möjliga.

b) Inte väntevärdesriktig. Låt $\eta = \min(\xi_1, \dots, \xi_n)$. Eftersom varje $\xi_i \geq \theta$ är $\eta \geq \theta$, och $\eta > \theta$ med sannolikhet 1. En variabel som alltid ligger över θ kan inte ha väntevärdet θ , så $E(\eta) > \theta$: skattningen underskattar systematiskt avståndet till ändpunkten.

Bevis. För $\theta < y < 0$ är $P(\xi_i > y) = \frac{0-y}{0-\theta} = \frac{y}{\theta}$. Oberoendet ger

$$P(\eta > y) = \left(\frac{y}{\theta}\right)^n$$

Eftersom η ligger i $[\theta, 0]$ är

$$E(\eta) = \theta + \int_{\theta}^0 \left(\frac{y}{\theta}\right)^n dy = \theta - \frac{\theta}{n+1} = \boxed{\frac{n}{n+1} \theta} > \theta$$

där sista olikheten gäller eftersom $\theta < 0$. Skattningen är alltså inte väntevärdesriktig, men $E(\eta) \rightarrow \theta$ när $n \rightarrow \infty$. Den korregerade skattningen $\frac{n+1}{n} \min(x_1, \dots, x_n)$ är väntevärdesriktig.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

15. Del 2: överlappar två konfidensintervall?

Antag att du har ett stickprov $\xi_1, \xi_2, \dots, \xi_n$ från $N(\mu_1, \sigma_1)$ och ett annat stickprov $\eta_1, \eta_2, \dots, \eta_n$ från $N(\mu_2, \sigma_2)$ där $\sigma_1 = \sigma_2 = 1$. De två stickproven är oberoende. För att testa $H_0: \mu_1 = \mu_2$ mot $H_1: \mu_1 \neq \mu_2$ skulle man kunna beräkna konfidensintervallen $I_1 = [\bar{\xi} - 1,96/\sqrt{n}, \bar{\xi} + 1,96/\sqrt{n}]$ och $I_2 = [\bar{\eta} - 1,96/\sqrt{n}, \bar{\eta} + 1,96/\sqrt{n}]$.

Vad är sannolikheten att dessa intervall överlappar varandra då H_0 är **falsk** och $\mu_1 = 4,0, \mu_2 = 4,4$?

Totalpoäng: 10

 Ledtråd 1

Två **konfidenstervall** med samma halva bredd c överlappar precis när medelvärdena ligger högst $2c$ från varandra.

 Ledtråd 2

Studera $\bar{\xi} - \bar{\eta}$. Enligt **Medelvärdets fördelning** och **Summor av normalfördelade stokastiska variabler** är $\bar{\xi} - \bar{\eta} \in N(-0,4, \sqrt{2/n})$.

 Ledtråd 3

Sök $P(|\bar{\xi} - \bar{\eta}| \leq 2 \cdot 1,96/\sqrt{n})$ och standardisera båda gränserna. Svaret beror på n .

 Lösning

Modell. $\bar{\xi} \in N(4,0; 1/\sqrt{n})$ och $\bar{\eta} \in N(4,4; 1/\sqrt{n})$ är oberoende, så

$$\bar{\xi} - \bar{\eta} \in N(-0,4, \sqrt{\frac{2}{n}})$$

Villkor för överlapp. Båda intervallen har halva bredden $c = 1,96/\sqrt{n}$. De överlappar om och endast om mittpunkterna ligger högst $2c$ från varandra:

$$|\bar{\xi} - \bar{\eta}| \leq \frac{3,92}{\sqrt{n}}$$

Räkning. Standardisera med $\sqrt{2/n}$:

$$P(\text{överlapp}) = \Phi\left(\frac{3,92/\sqrt{n} + 0,4}{\sqrt{2/n}}\right) - \Phi\left(\frac{-3,92/\sqrt{n} + 0,4}{\sqrt{2/n}}\right)$$

$$P(\text{överbäpp}) = \Phi\left(\frac{3,92}{\sqrt{2}} + 0,4\sqrt{\frac{n}{2}}\right) - \Phi\left(-\frac{3,92}{\sqrt{2}} + 0,4\sqrt{\frac{n}{2}}\right) = \Phi(2,772 + 0,283\sqrt{n}) - \Phi(-2,772 + 0,283\sqrt{n})$$

Några värden:

n	10	25	50	100
$P(\text{överbäpp})$	0,970	0,913	0,780	0,477

Slutsats. Om man förkastar H_0 bara när intervallen *inte* överbäppar blir testet mycket svagt: med $n = 25$ missar man skillnaden 0,4 i överbä 90 % av fallen. Överbäpp betyder alltså inte att $\mu_1 = \mu_2$. Det rätta testet bygger på ett intervall för $\mu_1 - \mu_2$ med halva bredden $1,96\sqrt{2/n}$, som är smalare än $2 \cdot 1,96/\sqrt{n}$.

I uppgiftstexten står mothypotesen som " $H_0 : \mu_1 \neq \mu_2$ "; det är ett skrivfel för H_1 .

På räknaren. För $n = 25$: `normCdf(-3.92/5, 3.92/5, -0.4, sqrt(2/25))`.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överbä.

Se även

- S0008M
- Betingad sannolikhet
- Oberoende händelser
- Linjära funktioner av stokastiska variabler
- Hypergeometrisk fördelning
- Likformig fördelning
- Centrala gränsvärdessatsen
- Signifikansnivå och styrka
- Två stickprov
- Stickprov i par
- Direktmetoden
- Tvådimensionell normalfördelning
- Punktskattning
- Konfidensintervall