

S0008M-2024-10-26

19 okt. 2026, 16 min läsning

UTKAST

#tenta #sannolikhetslära #statistik

Kurs: S0008M — Sannolikhetslära och statistik **Datum:** 2024-10-26 **Källa:** originaltentamen ·
lärarens lösningar

Översikt

Del 1 (uppgift 1–12, bara svar):

- Bayes sats för kasserade detaljer (1) och standardavvikelse för böter (2)
- ffg-fördelning (3) och μ, σ ur två andelar i en normalfördelning (4)
- differens av normalfördelade variabler (5) och exponentialfördelningens glömska (6)
- CGS för en summa av väntetider (7) och stickprovsstorlek för given styrka (8)
- konfidensintervall med stickprov i par (9) och antal intervall som täcker (10)
- likformig fördelning på kvadraten (11) och maximum av tre oberoende variabler (12)

Del 2 (uppgift 13–15):

- icke-heltalsdelen av en summa av två likformiga variabler (13)
- linjära kombinationer som tvådimensionell normalfördelning (14)
- väntevärdesriktighet och fördelning för en andelsskattning (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, kontrollerad mot lärarens facit.

1. Är den kasserade detaljen defekt?

Vid tillverkning av detaljer till en maskin är felfrekvensen 1 %. För att begränsa antalet reklamationer beslutar man att alla detaljer ska passera en kontroll. Vid denna kontroll kasseras felaktiga detaljer med sannolikheten 0,93, och felfria kasseras med sannolikheten 0,01. Du tar en detalj ur högen med kasserade detaljer. Hur stor är sannolikheten att den detaljen är defekt? Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Du vet att detaljen är kasserad och frågar efter orsaken: **Bayes sats**.

Ledtråd 2

Låt D = defekt och K = kasserad. $P(D \mid K) = \frac{P(K \mid D)P(D)}{P(K)}$.

Ledtråd 3

Nämnaren fås ur **lagen om total sannolikhet**: $P(K) = P(K \mid D)P(D) + P(K \mid D^*)P(D^*)$.

Lösning

Låt D vara händelsen att detaljen är defekt och K att den kasseras. Lagen om total sannolikhet ger

$$P(K) = 0,93 \cdot 0,01 + 0,01 \cdot 0,99 = 0,0093 + 0,0099 = 0,0192$$

Bayes sats ger sedan

$$P(D \mid K) = \frac{0,93 \cdot 0,01}{0,0192} = 0,484375 \approx \boxed{48,44 \%}$$

Kontrollerat mot facit.

2. Standardavvikelse för böterna

Vid ett konstruktionsarbete kan tiden med vilken arbetet försenas betraktas som en diskret stokastisk variabel ξ med följande sannolikhetsfunktion:

$$P(\xi = 0) = 0,5, \quad P(\xi = 1) = 0,3, \quad P(\xi = 2) = 0,2.$$

Om arbetet försenas får konstruktionsfirman böta, dels en fast kostnad på 10 000 kr, dels 15 000 kr per försenad dag. Bestäm **standardavvikelsen** för böterna. Svara med en decimal.

Totalpoäng: 2

Ledtråd 1

Böterna är en ny diskret stokastisk variabel. Räkna **väntevärdet** och **variansen** direkt ur dess sannolikhetsfunktion.

Ledtråd 2

Den fasta kostnaden betalas bara om arbetet försenas. Därför är böterna **inte** $10\,000 + 15\,000 \xi$.

Ledtråd 3

Böterna η antar värdena 0, 25 000 och 40 000 med sannolikheterna 0,5, 0,3 och 0,2.

Lösning

Låt η vara böterna. Utan försening blir det inga böter, så

ξ	0	1	2
η (kr)	0	25 000	40 000
P	0,5	0,3	0,2

$$E(\eta) = 0 \cdot 0,5 + 25\,000 \cdot 0,3 + 40\,000 \cdot 0,2 = 15\,500$$

$$V(\eta) = (0 - 15\,500)^2 \cdot 0,5 + (25\,000 - 15\,500)^2 \cdot 0,3 + (40\,000 - 15\,500)^2 \cdot 0,2 = 267\,250\,000$$

$$D(\eta) = \sqrt{267\,250\,000} \approx \boxed{16\,347,8 \text{ kr}}$$

Fällan är att skriva $\eta = 10\,000 + 15\,000 \xi$ och använda **satsen om linjära funktioner**. Det ger fel svar, $15\,000 \cdot \sqrt{0,61} \approx 11\,715,4$, eftersom den fasta kostnaden inte betalas när $\xi = 0$.

Kontrollerat mot facit.

3. Kast tills första klave

Du kastar ett mynt tills du får klave. Låt ξ stå för antalet kast som du behöver göra totalt för att få den första klaven (inklusive kastet som ger klave). De möjliga värdena på ξ är alltså heltalen 1, 2, 3, ...

Beräkna sannolikheten $P(3 \leq \xi \leq 5)$. Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Antal försök till och med första lyckade: **ffg-fördelning**.

Ledtråd 2

$\xi \in \text{ffg}(0,5)$, så $P(\xi = k) = 0,5 \cdot 0,5^{k-1} = 0,5^k$.

Ledtråd 3

Summera $k = 3, 4, 5$, eller använd $P(\xi > k) = 0,5^k$ (de k första kasten ger krona).

Lösning

$\xi \in \text{ffg}(0,5)$ ger $P(\xi = k) = 0,5^k$, så

$$P(3 \leq \xi \leq 5) = 0,5^3 + 0,5^4 + 0,5^5 = 0,125 + 0,0625 + 0,03125 = 0,21875$$

Samma sak med komplement: $P(\xi > k) = 0,5^k$ ger $P(3 \leq \xi \leq 5) = P(\xi > 2) - P(\xi > 5) = 0,5^2 - 0,5^5$.

21,88 %

Kontrollerat mot facit.

4. μ och σ ur två andelar

Diametern för axlar som tillverkas kan betraktas som en stokastisk variabel ξ . Vid kvalitetskontroll sorterar man bort de axlar som är tjockare än 1,07 mm eller smalare än 0,91 mm. Man har funnit att $P(\xi > 1,07) = 0,09$ och $P(\xi < 0,91) = 0,03$.

Antag att $\xi \in N(\mu, \sigma)$. Bestäm μ och σ och ange värdet på σ . Svara med tre decimaler.

Totalpoäng: 2

Ledtråd 1

Två kända andelar ger två **kvantiler** i en **normalfördelning**: två ekvationer, två okända.

Ledtråd 2

Standardisera: $\frac{1,07 - \mu}{\sigma} = \lambda_{0,09}$ och $\frac{0,91 - \mu}{\sigma} = -\lambda_{0,03}$.

Ledtråd 3

Subtrahera ekvationerna: $\frac{0,16}{\sigma} = \lambda_{0,09} + \lambda_{0,03}$.

 Lösning

Standardisering ger

$$\frac{1,07 - \mu}{\sigma} = \lambda_{0,09} \approx 1,3408, \quad \frac{0,91 - \mu}{\sigma} = -\lambda_{0,03} \approx -1,8808$$

Subtraktion ger

$$\sigma = \frac{1,07 - 0,91}{1,3408 + 1,8808} = \frac{0,16}{3,2216} \approx 0,0497 \approx \boxed{0,050}$$

och $\mu = 1,07 - 1,3408 \cdot 0,0497 \approx 1,003$ mm.

På räknaren. `invNorm(0.91)` och `invNorm(0.97)` ger kvantilerna.

Kontrollerat mot facit.

5. $P(\xi_1 > 1 + \xi_2)$

Antag att ξ_1 och ξ_2 är oberoende och $N(0, 1)$ -fördelade. Beräkna sannolikheten $P(\xi_1 > 1 + \xi_2)$. Svara i procent med två decimaler.

Totalpoäng: 2

 Ledtråd 1

Samla variablerna på ena sidan: det blir en **linjärkombination** av normalfördelade variabler.

 Ledtråd 2

$$\xi_1 - \xi_2 \in N(0, \sqrt{1+1}) = N(0, \sqrt{2}).$$

 Ledtråd 3

$$P(\xi_1 - \xi_2 > 1) = 1 - \Phi(1/\sqrt{2}).$$

 Lösning

$P(\xi_1 > 1 + \xi_2) = P(\xi_1 - \xi_2 > 1)$, och $\xi_1 - \xi_2 \in N(0, \sqrt{2})$ eftersom varianserna adderas. Alltså

$$P(\xi_1 - \xi_2 > 1) = 1 - \Phi\left(\frac{1}{\sqrt{2}}\right) = 1 - \Phi(0,7071) \approx \boxed{23,98 \%}$$

På räknaren. `normCdf(1,∞,0,√(2))`.

Kontrollerat mot facit. Läraren godtar också $1 - \Phi(0,71) \approx 23,89 \%$ med avrundat tabellvärde.

6. Betingad sannolikhet i exponentialfördelningen

Antag att $\xi \in \text{Exp}(1)$. Beräkna den betingade sannolikheten $P(2 \leq \xi \leq 3 \mid \xi \geq 1)$. Svara i procent med två decimaler.

Totalpoäng: 2

 Ledtråd 1

Exponentialfördelningen saknar minne.

 Ledtråd 2

Givet $\xi \geq 1$ är $\xi - 1$ åter $\text{Exp}(1)$. Därför är $P(2 \leq \xi \leq 3 \mid \xi \geq 1) = P(1 \leq \xi \leq 2)$.

 Ledtråd 3

$F(x) = 1 - e^{-x}$, så $P(1 \leq \xi \leq 2) = F(2) - F(1)$.

 Lösning

Glömskan ger att intervallet kan flyttas ett steg åt vänster:

$$P(2 \leq \xi \leq 3 \mid \xi \geq 1) = P(1 \leq \xi \leq 2) = F(2) - F(1) = e^{-1} - e^{-2} \approx \boxed{23,25 \%}$$

Direkt med definitionen blir det samma: $\frac{e^{-2} - e^{-3}}{e^{-1}} = e^{-1} - e^{-2}$.

Kontrollerat mot facit.

7. Femtio fiskar innan hemgång

Thomas är ute och fiskar. Hans väntetider mellan napp (då en fisk fastnar på kroken) antas vara oberoende och exponentialfördelade med väntevärde 7,5 minuter. Thomas bestämmer sig för att inte gå hem förrän hon fått 50 fiskar.

Vad är sannolikheten för att det tar mer än 8 timmar (dvs mer än 480 minuter) innan Thomas får gå hem? Svara i procent med en decimal.

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

Totalpoäng: 2

 Ledtråd 1

Den totala tiden är en summa av 50 oberoende väntetider: Centrala gränsvärdessatsen.

 Ledtråd 2

En exponentialfördelad väntetid med väntevärde 7,5 har också standardavvikelse 7,5.

 Ledtråd 3

Summan är $\approx N(50 \cdot 7,5, 7,5\sqrt{50}) = N(375; 53,03)$.

 Lösning

Låt $\eta = \xi_1 + \dots + \xi_{50}$ vara tiden till den 50:e fisken. Varje ξ_i har $E(\xi_i) = D(\xi_i) = 7,5$, så

$$E(\eta) = 50 \cdot 7,5 = 375, \quad D(\eta) = 7,5\sqrt{50} \approx 53,03$$

CGS ger

$$P(\eta > 480) \approx 1 - \Phi\left(\frac{480 - 375}{53,03}\right) = 1 - \Phi(1,980) \approx \boxed{2,4 \%}$$

På räknaren. `normCdf(480,∞,375,7.5*√(50))`.

Det exakta värdet (summan är gammafördelad) är ungefär 3,1 %. Normalapproximationen underskattar svansen något eftersom summan är skev åt höger.

Kontrollerat mot facit. Lärarens facit skriver parametrarna som $N(50 \cdot 1/7,5, \sqrt{50} \cdot 1/7,5)$ men anger värdena 375 och 53,03, alltså $50 \cdot 7,5$ och $\sqrt{50} \cdot 7,5$; det är ett skrivfel och svaret 2,4 % stämmer.

8. Stickprovsstorlek för 90 % styrka

Antag att $\xi_1, \xi_2, \xi_3, \dots, \xi_n$ är ett stickprov av storlek n från $N(\mu; 0,55)$ och att man vill testa $H_0 : \mu = 10$ mot $H_1 : \mu = 10,2$ på 5 % signifikansnivå. Hur stort måste n vara för att det test som har 5 % signifikansnivå ska få en styrka som är minst 90 %?

Svara med det minsta värdet på n (ett heltal) för vilket gäller att testet med 5 % signifikansnivå har en styrka som är minst 90 %.

Totalpoäng: 2

 Ledtråd 1

Signifikansnivå och styrka för enkla hypoteser: förkasta H_0 om $\bar{\xi} > k$.

 Ledtråd 2

$\bar{\xi} \in N(\mu; 0,55/\sqrt{n})$. Nivån ger $\frac{k-10}{0,55/\sqrt{n}} = \lambda_{0,05}$ och styrkan ger $\frac{k-10,2}{0,55/\sqrt{n}} = -\lambda_{0,10}$.

 Ledtråd 3

Subtrahera: $\frac{0,2}{0,55/\sqrt{n}} = \lambda_{0,05} + \lambda_{0,10}$. Lös ut n och avrunda uppåt.

 Lösning

Testet förkastar H_0 om $\bar{\xi} > k$, där $\bar{\xi} \in N(\mu; 0,55/\sqrt{n})$ (**Medelvärdets fördelning**). Kraven blir

$$\frac{k-10}{0,55/\sqrt{n}} = \lambda_{0,05} = 1,6449, \quad \frac{k-10,2}{0,55/\sqrt{n}} = -\lambda_{0,10} = -1,2816$$

Subtraktion ger

$$\sqrt{n} = \frac{(1,6449 + 1,2816) \cdot 0,55}{0,2} \approx 8,048, \quad n \approx 64,76$$

Med $n = 64$ blir styrkan 89,7 % och med $n = 65$ blir den 90,1 %:

$$\boxed{n = 65}$$

Kontrollerat mot facit.

9. Brinntid för facklor: stickprov i par

Vid en jämförelse av brinntiden hos ljusfacklor av två olika fabrikat, A och B, har man mätt brinntiden under 6 olika dagar. Varje dag valde man slumpmässigt en fackla av varje fabrikat. De yttre förhållandena såsom vind, luftfuktighet och temperatur varierade mellan de olika dagarna. Resultatet (enhet: minuter) blev följande för A:

209,4	552,4	354,4	395,3	284,4	448,6
-------	-------	-------	-------	-------	-------

För B blev resultatet (enhet: minuter) följande:

198,1	529,5	351,4	423,1	311,2	428,6
-------	-------	-------	-------	-------	-------

Beräkna ett 95 % konfidensintervall för den genomsnittliga skillnaden i brinntid mellan facklor av typ A och facklor av typ B, definierad som genomsnittlig brinntid för facklor av typ A minus genomsnittlig brinntid för facklor av typ B. Antag att en facklas brinntid på en given dag kan beskrivas med en normalfördelad stokastisk variabel. Ange intervallets övre gräns med två decimaler.

Totalpoäng: 2

Ledtråd 1

Förhållandena varierar mellan dagarna men är lika för paret samma dag: **Stickprov i par**.

Ledtråd 2

Bilda differenserna $z_i = x_i - y_i$ och gör ett [[t-fördelningen|t-intervall]] för deras väntevärde: $\bar{z} \pm t_{0,025}(5) s_z / \sqrt{6}$.

Ledtråd 3

Differenserna är 11,3, 22,9, 3,0, -27,8, -26,8 och 20,0.

Lösning

Dagarna ger olika förhållanden, så facklorna jämförs parvis. Differenserna $z_i = x_i - y_i$ är

11,3	22,9	3,0	-27,8	-26,8	20,0
------	------	-----	-------	-------	------

$$\bar{z} = 0,4333, \quad s_z = 22,588, \quad t_{0,025}(5) = 2,5706$$

Konfidensintervallet blir

$$I_{\Delta} = 0,4333 \pm 2,5706 \cdot \frac{22,588}{\sqrt{6}} = 0,43 \pm 23,70 = [-23,27; 24,14]$$

Övre gräns: 24,14 minuter.

På räknaren. `invT(0.975,5)` ger 2,5706. Eller lägg differenserna i en lista och kör `tInterval` med konfidensgrad 0,95.

Kontrollerat mot facit.

10. Hur många av 14 intervall täcker?

En forskare har konstruerat konfidensintervall för 14 olika okända konstanter i 14 olika fördelningar. Varje intervall har konfidensgrad 90 % och de olika intervallen härrör från av varandra oberoende mätserier. Standardavvikelserna i de 14 fördelningarna är okända.

Vad är sannolikheten att minst 11 av de 14 intervallen täcker den konstant för vilken intervallet konstruerats? Svara i procent med två decimaler.

Totalpoäng: 2

Ledtråd 1

Varje intervall täcker eller inte, oberoende av de andra: **Binomialfördelning**.

Ledtråd 2

Konfidensgraden är sannolikheten att ett intervall täcker, så antalet som täcker är $\xi \in \text{Bin}(14; 0,9)$.

Ledtråd 3

$P(\xi \geq 11) = 1 - P(\xi \leq 10)$. Att standardavvikelserna är okända spelar ingen roll.

 Lösning

Antalet intervall som täcker sin konstant är $\xi \in \text{Bin}(14; 0,9)$. Alltså

$$P(\xi \geq 11) = 1 - P(\xi \leq 10) = \sum_{k=11}^{14} \binom{14}{k} 0,9^k 0,1^{14-k} \approx \boxed{95,59 \%}$$

På räknaren. `binomCdf(14,0.9,11,14)`.

Kontrollerat mot facit.

11. Punkt inom enhetscirkeln

Slumpvariablerna ξ och η är oberoende och $R(0, 1)$ -fördelade. Beräkna sannolikheten $P(\xi^2 + \eta^2 \leq 1)$. Svara i procent med två decimaler.

Totalpoäng: 3

 Ledtråd 1

(ξ, η) är likformigt fördelad på enhetskvadraten (Flerdimensionella fördelningar). Sannolikhet blir area.

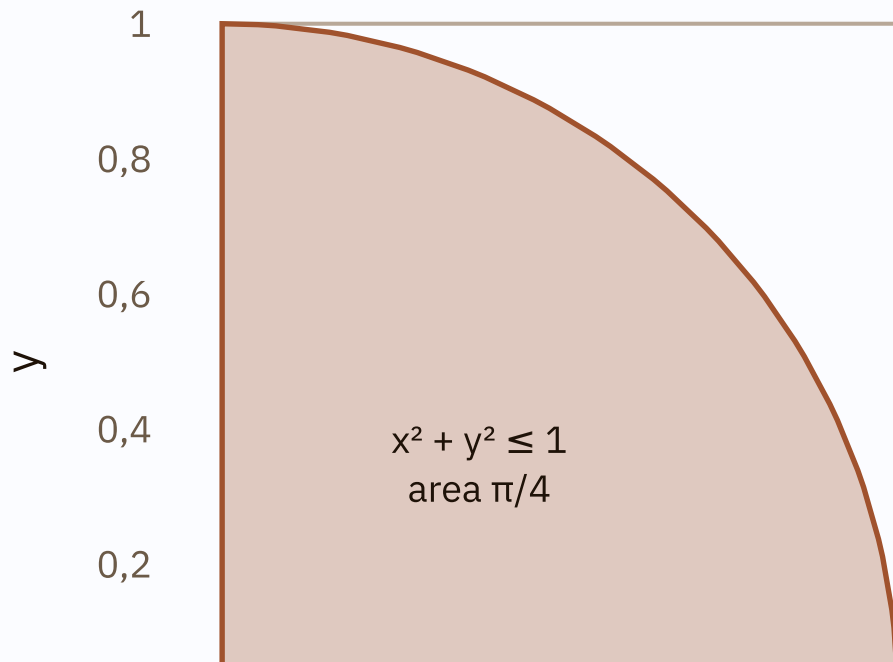
 Ledtråd 2

$f(x, y) = 1$ på $0 < x, y < 1$. Sannolikheten är arean av den del av området $x^2 + y^2 \leq 1$ som ligger i kvadraten.

 Ledtråd 3

Den delen är en kvarts cirkelskiva med radie 1.

 Lösning



Eftersom ξ och η är **oberoende** är $f(x, y) = 1 \cdot 1 = 1$ på kvadraten $K = \{0 < x, y < 1\}$, som har arean 1. Villkoret $\xi^2 + \eta^2 \leq 1$ betyder att punkten ligger i enhetscirkeln, och dess del i K är en kvartscirkel:

$$P(\xi^2 + \eta^2 \leq 1) = \frac{\pi \cdot 1^2/4}{1} = \frac{\pi}{4} \approx \boxed{78,54 \%}$$

Kontrollerat mot facit.

12. Maximum av tre olika fördelningar

Antag att ξ_1 , ξ_2 och ξ_3 är oberoende, där $\xi_1 \in R(0, 2)$, $\xi_2 \in N(0, 1)$, $\xi_3 \in \text{Exp}(1)$. Beräkna sannolikheten $P(\max(\xi_1, \xi_2, \xi_3) > 1)$. Svara i procent med två decimaler.

Totalpoäng: 3

Ledtråd 1

Maximum: gå via komplementet.

Ledtråd 2

$\max \leq 1$ betyder att **alla tre** är ≤ 1 . Oberoende ger en produkt.

Ledtråd 3

$$P(\xi_1 \leq 1) = 0,5, P(\xi_2 \leq 1) = \Phi(1) \text{ och } P(\xi_3 \leq 1) = 1 - e^{-1}.$$

Lösning

$$P(\max(\xi_1, \xi_2, \xi_3) > 1) = 1 - P(\xi_1 \leq 1)P(\xi_2 \leq 1)P(\xi_3 \leq 1)$$

med

$$P(\xi_1 \leq 1) = \frac{1}{2}, \quad P(\xi_2 \leq 1) = \Phi(1) = 0,8413, \quad P(\xi_3 \leq 1) = 1 - e^{-1} = 0,6321$$

Produkten är $0,5 \cdot 0,8413 \cdot 0,6321 = 0,2659$, så

$$P(\max > 1) = 1 - 0,2659 \approx \boxed{73,41 \%}$$

Kontrollerat mot facit.

13. Del 2: icke-heltalsdelen av $\xi_1 + \xi_2$

Uppgift 13–15 utgör del 2 av tentan. Om du har gjort Lab 4 i kursen under LP1 så skriv “Lab 4 inlämnad” på lösningsbladet för denna uppgift.

Låt ξ_1 och ξ_2 vara oberoende och $R(0, 1)$ -fördelade och låt $\xi = \xi_1 + \xi_2$. Låt $\zeta = [\xi]$ vara icke-heltalsdelen av den stokastiska variabeln ξ (det gäller t.ex. att $[2,89] = 0,89$ och $[4,07] = 0,07$).

Visa att ζ är $R(0, 1)$ -fördelad.

Tips: Bestäm fördelningsfunktionen för ζ .

Totalpoäng: 10

Ledtråd 1

Funktion av en stokastisk variabel: räkna ut $F_\zeta(z) = P(\zeta \leq z)$ via fördelningsfunktionen för ξ .

Ledtråd 2

Summan av två oberoende $R(0, 1)$ har triangelthet på $(0, 2)$: $F_\xi(x) = x^2/2$ för $0 < x < 1$ och $F_\xi(x) = 1 - (2 - x)^2/2$ för $1 \leq x < 2$.

Ledtråd 3

För $0 < z < 1$ är $\zeta \leq z$ samma sak som $0 \leq \xi \leq z$ eller $1 \leq \xi \leq 1 + z$.

Lösning

Modell. $\xi = \xi_1 + \xi_2$ där ξ_1, ξ_2 är oberoende och $R(0, 1)$. Fördelningsfunktionen för summan är

$$F_\xi(x) = \begin{cases} x^2/2, & 0 < x < 1, \\ 1 - (2 - x)^2/2, & 1 \leq x < 2. \end{cases}$$

Den fås ur arean av den del av enhetskvadraten där $x_1 + x_2 \leq x$ (Vännman, exempel 11.4).

Metod. Eftersom $0 < \zeta < 1$ är $F_\zeta(z) = 0$ för $z \leq 0$ och $F_\zeta(z) = 1$ för $z \geq 1$. Låt $0 < z < 1$. Eftersom $0 < \xi < 2$ gäller

$$\{\zeta \leq z\} = \{0 \leq \xi \leq z\} \cup \{1 \leq \xi \leq 1+z\}$$

och händelserna är disjunkta.

Räkning.

$$\begin{aligned} F_{\zeta}(z) &= F_{\xi}(z) + F_{\xi}(1+z) - F_{\xi}(1) \\ &= \frac{z^2}{2} + \left(1 - \frac{(1-z)^2}{2}\right) - \frac{1}{2} = \frac{z^2}{2} + \frac{1}{2} - \frac{1-2z+z^2}{2} = z \end{aligned}$$

Slutsats.

$$F_{\zeta}(z) = z, \quad 0 < z < 1$$

Det är fördelningsfunktionen för $R(0, 1)$, så ζ är $R(0, 1)$ -fördelad.

Kontrollerat mot facit.

14. Del 2: linjära kombinationer av två normalfördelade variabler

Antag att ζ_1 och ζ_2 är oberoende och $N(0, 1)$ -fördelade.

a) Låt $\xi = 1 + \zeta_1 + 2\zeta_2$, $\eta = 3\zeta_1 + \zeta_2$. Då har (ξ, η) en tvådimensionell kontinuerlig fördelning. Vilken? Ange typ av fördelning och relevanta parametrar. (8 p)

b) Låt $\zeta_3 = -\zeta_1$. Förklara varför (ζ_1, ζ_3) inte har en tvådimensionell kontinuerlig fördelning. (2 p)

Totalpoäng: 10

Ledtråd 1

Tvådimensionell normalfördelning: (ξ, η) är tvådimensionellt normalfördelad om varje linjärkombination $c_1\xi + c_2\eta$ är normalfördelad.

Ledtråd 2

Parametrarna är $\mu_\xi, \mu_\eta, \sigma_\xi, \sigma_\eta$ och ρ . Räkna **kovariansen** genom att multiplicera ut $E((\zeta_1 + 2\zeta_2)(3\zeta_1 + \zeta_2))$.

Ledtråd 3

$E(\zeta_i^2) = 1$ och $E(\zeta_1\zeta_2) = 0$. I b): vad är $\zeta_1 + \zeta_3$?

Lösning

a) Modell. En linjärkombination $c_1\xi + c_2\eta = c_1 + (c_1 + 3c_2)\zeta_1 + (2c_1 + c_2)\zeta_2$ är en linjärkombination av oberoende normalfördelade variabler och därför normalfördelad. Alltså är (ξ, η) tvådimensionellt normalfördelad.

Räkning.

$$E(\xi) = 1, \quad E(\eta) = 0$$

$$V(\xi) = 1^2 + 2^2 = 5, \quad V(\eta) = 3^2 + 1^2 = 10$$

$$C(\xi, \eta) = E((\zeta_1 + 2\zeta_2)(3\zeta_1 + \zeta_2)) = E(3\zeta_1^2 + 7\zeta_1\zeta_2 + 2\zeta_2^2) = 3 + 0 + 2 = 5$$

$$\rho = \frac{5}{\sqrt{5} \cdot \sqrt{10}} = \frac{1}{\sqrt{2}} \approx 0,707$$

Slutsats.

$$(\xi, \eta) \text{ tvådimensionellt normalfördelad med } \mu_\xi = 1, \mu_\eta = 0, \sigma_\xi = \sqrt{5}, \sigma_\eta = \sqrt{10}, \rho = \frac{1}{\sqrt{2}}$$

b) $\zeta_1 + \zeta_3 = \zeta_1 - \zeta_1 = 0$ är konstant och alltså inte normalfördelad. Då kan (ζ_1, ζ_3) inte vara tvådimensionellt normalfördelad. Man kan också se att **korrelationen** är $\rho = -1$, och då är tätheten för tvådimensionell normalfördelning inte definierad.

Mer allmänt ligger (ζ_1, ζ_3) alltid på linjen $y = -x$. En linje har area noll, så ingen tvådimensionell frekvensfunktion kan ge den sannolikheten 1. Fördelningen är därför inte kontinuerlig.

Kontrollerat mot facit.

15. Del 2: andelsskattningen p^*

Vid tillverkning av en viss typ av komponenter blir en viss okänd andel p av komponenterna felaktiga. Man tar ut n komponenter på måfå och skattar p med $p^* =$ andelen felaktiga komponenter, dvs antalet felaktiga komponenter delat med n .

(a) Är p^* väntevärdesriktig? Motivera. (4 p)

(b) Vad blir den approximativa fördelningen för p^* om n är stort? Motivera. (6 p)

Totalpoäng: 10

Ledtråd 1

Punktskattning: en skattning är väntevärdesriktig om $E(p^*) = p$.

Ledtråd 2

Antalet felaktiga är $\xi \in \text{Bin}(n; p)$ och $p^* = \xi/n$.

Ledtråd 3

Använd $E(\xi) = np$ och $V(\xi) = np(1 - p)$. För stort n : **Centrala gränsvärdessatsen**.

Lösning

Modell. Låt ξ vara antalet felaktiga bland de n komponenterna. Komponenterna väljs på måfå, så $\xi \in \text{Bin}(n; p)$ och $p^* = \xi/n$.

(a) Enligt **satsen om linjära funktioner** är

$$E(p^*) = \frac{1}{n}E(\xi) = \frac{1}{n} \cdot np = p$$

Ja, p^* är väntevärdesriktig.

(b) Variansen är

$$V(p^*) = \frac{1}{n^2}V(\xi) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

ξ är en summa av n oberoende indikatorer (1 om komponenten är felaktig, annars 0). För stort n ger CGS att ξ , och därmed $p^* = \xi/n$, är approximativt normalfördelad (**Normalapproximation**):

$$p^* \approx N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Kontrollerat mot facit.

Se även

- [S0008M](#)
- [Bayes sats](#) — uppgift 1
- [ffg-fördelning](#) — uppgift 3
- [Exponentialfördelning](#) — uppgift 6
- [Centrala gränsvärdessatsen](#) — uppgift 7 och 15
- [Signifikansnivå och styrka](#) — uppgift 8
- [Stickprov i par](#) — uppgift 9
- [Max och min av stokastiska variabler](#) — uppgift 12
- [Icke-linjära funktioner av stokastiska variabler](#) — uppgift 13
- [Tvådimensionell normalfördelning](#) — uppgift 14
- [Punktskattning](#) — uppgift 15