

Tvådimensionell normalfördelning

19 okt. 2026, 9 min läsning

UTKAST

#matematik #sannolikhetslära #fördelning

Kurs: S0008M Förekunskaper: Flerdimensionella fördelningar, Korrelation, Normalfördelning

Den tvådimensionella normalfördelningen beskriver två normalfördelade variabler tillsammans och bestäms helt av de två väntevärdena, de två standardavvikelseerna och korrelationen. För just denna fördelning gäller att okorrelerade variabler också är oberoende.

1. Varför den behövs

Väljer vi två personer slumpmässigt och låter ξ_1 vara den enas längd och ξ_2 den andras vikt är de oberoende, och den simultana frekvensfunktionen är produkten av två **normalfördelningar**. Väljer vi i stället *en* person och mäter både längd och vikt är båda fortfarande normalfördelade var för sig, men de är inte oberoende. Vi behöver en modell där marginalerna är $N(\mu_1, \sigma_1)$ och $N(\mu_2, \sigma_2)$ men där variablerna får vara beroende.

2. Definitionen

Definition 1.20 (kompendiet). (ξ_1, ξ_2) har en **tvådimensionell normalfördelning** om den **simultana frekvensfunktionen** är

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} \right]\right)$$

där $\sigma_1, \sigma_2 > 0$ och $-1 < \rho < 1$.

De fem parametrarna betyder: $\xi_1 \in N(\mu_1, \sigma_1)$, $\xi_2 \in N(\mu_2, \sigma_2)$ och $\rho = \rho(\xi_1, \xi_2)$ är **korrelationen**.

Okorrelerade $\Rightarrow$ oberoende

Med $\rho = 0$ faller korstermen bort och $f(x_1, x_2)$ blir produkten av de två marginalerna. För tvådimensionellt normalfördelade variabler är okorrelerade alltså samma sak som oberoende — det gäller inte i allmänhet.

Nivåkurvorna är ellipser. Med $\rho > 0$ lutar de uppåt åt höger, och ju närmare $|\rho|$ är 1, desto smalare blir de.

3. Linjärkombinationer

Sats 1.21 (kompendiet). Om (ξ_1, ξ_2) har en tvådimensionell normalfördelning och c_1, c_2 är konstanter, inte båda noll, så är $c_1\xi_1 + c_2\xi_2$ (endimensionellt) normalfördelad.

Väntevärde och varians tar vi som vanligt, där **kovariansen** nu kommer med:

$$\xi_1 \pm \xi_2 \in N(\mu_1 \pm \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2 \pm 2\rho\sigma_1\sigma_2})$$

Frågor om summor och differenser förs därmed över till den vanliga normalfördelningen och Φ -tabellen.

4. Betingad fördelning

Sats 1.23 (kompendiet). Den **betingade fördelningen** för ξ_2 givet $\xi_1 = x_1$ är $N(\mu'_2, \sigma'_2)$ med

$$\mu'_2 = \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \quad \sigma'_2 = \sqrt{1 - \rho^2} \sigma_2$$

- **Väntevärdet** flyttas längs en linje i x_1 . Är $\rho > 0$ och $x_1 > \mu_1$ blir $\mu'_2 > \mu_2$: ett stort värde på ξ_1 talar för ett stort värde på ξ_2 .
- **Standardavvikelsen** minskar med faktorn $\sqrt{1 - \rho^2}$ och beror inte på x_1 . Information om ξ_1 minskar alltid osäkerheten om ξ_2 när variablerna är korrelerade.

Tvärsnittet genom ytan

Geometriskt är den betingade frekvensfunktionen ett lodrätt tvärsnitt genom $f(x_1, x_2)$ vid x_1 , normerat så att arean blir 1. Tvärsnittet är en normalkurva med toppen på den streckade linjen i figuren i exemplet nedan.

5. Lösta lektionsuppgifter

Övningarna märkta *Kompendiet* kommer ur kompendiet om flerdimensionella fördelningar. Det saknar facit, så de är lösta två gånger oberoende av varandra.

≡ Lektion 18, övning 2 – Bertils tiokamp

De två första grenarna i tiokamp är 100 meter och längdhopp. Tiokamparen Bertil förväntar sig 950 respektive 1025 poäng. Resultaten (ξ_1, ξ_2) beskrivs med en tvådimensionell normalfördelning med väntevärden 950 och 1025, standardavvikelser $\sigma_1 = 120$ och $\sigma_2 = 190$ och korrelation $\rho_{12} = 0,55$.

a) Beräkna sannolikheten att det totala resultatet efter två grenar är minst 2100.

b) Beräkna sannolikheten att resultatet i längd blir bättre än resultatet på 100 m.

Resultatet på 100 meter blir ett personbästa på 11,34, vilket ger 1083 poäng.

c) Bertils personbästa i längd motsvarar 1240 poäng. Vad är den betingade sannolikheten att det blir personbästa även i längd?

d) Vad är det betingade väntevärdet av det totala resultatet efter två grenar?

e) I vilken mening ger svaret i d) den bästa prognosen för det totala resultatet, givet resultatet i första grenen?

Lösning

a) Enligt kompendiets sats 1.21 är summan normalfördelad:

$$\xi_1 + \xi_2 \in N(1975, \sqrt{120^2 + 190^2 + 2 \cdot 0,55 \cdot 120 \cdot 190}) = N(1975; 274,9)$$

$$P(\xi_1 + \xi_2 \geq 2100) = 1 - \Phi\left(\frac{2100 - 1975}{274,9}\right) = 1 - \Phi(0,455) \approx \boxed{0,325}$$

b) "Längd bättre än 100 m" är $\xi_2 - \xi_1 > 0$, och differensen är också normalfördelad:

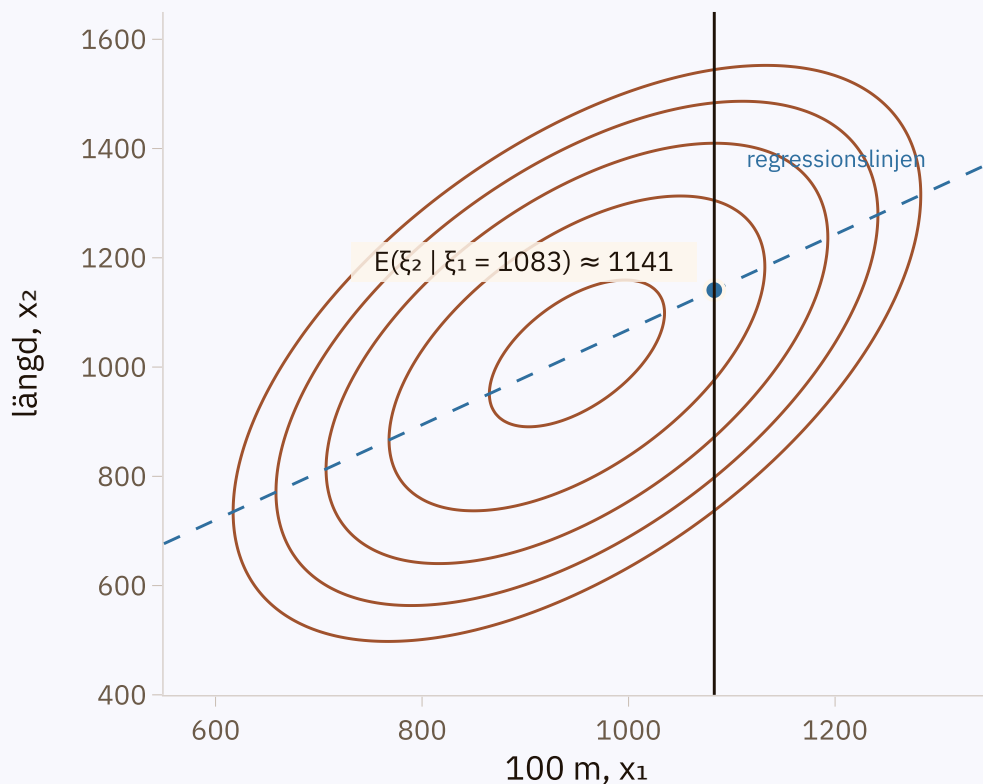
$$\xi_2 - \xi_1 \in N(75, \sqrt{120^2 + 190^2 - 2 \cdot 0,55 \cdot 120 \cdot 190}) = N(75; 159,4)$$

$$P(\xi_2 - \xi_1 > 0) = 1 - \Phi\left(\frac{0 - 75}{159,4}\right) = \Phi(0,470) \approx \boxed{0,681}$$

c) Kompendiets sats 1.23 med $x_1 = 1083$:

$$\mu'_2 = 1025 + 0,55 \cdot \frac{190}{120} \cdot (1083 - 950) \approx 1140,8, \quad \sigma'_2 = \sqrt{1 - 0,55^2} \cdot 190 \approx 158,7$$

$$P(\xi_2 \geq 1240 \mid \xi_1 = 1083) = 1 - \Phi\left(\frac{1240 - 1140,8}{158,7}\right) = 1 - \Phi(0,625) \approx \boxed{0,266}$$



Figuren visar nivåkurvorna, linjen $x_1 = 1083$ och det betingade väntevärdet på den streckade linjen $\mu'_2 = \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$.

d) 100-metersresultatet är känt, så bara längden är slumpmässig:

$$E(\xi_1 + \xi_2 \mid \xi_1 = 1083) = 1083 + \mu'_2 \approx \boxed{2223,8}$$

e) Det betingade väntevärdet är den prognos som minimerar det förväntade **kvadratiske** felet, givet informationen $\xi_1 = 1083$ – bästa prognosen i minsta-kvadrat-mening.

På räknaren. a) `normCdf(2100, ∞, 1975, 274.9)`, b) `normCdf(0, ∞, 75, 159.4)`, c) `normCdf(1240, ∞, 1140.8, 158.7)`. Räknaren ger facits fyra decimaler direkt.

Facit: a) 0,3246; b) 0,6809; c) 0,2659; d) 2223,8; e) bästa prognosen i minsta-kvadrat-mening.

☰ Kompendiet, övning 1.18 – Majas och Joels utgifter

Maja och Joel betraktar sina månatliga utgifter för kursmaterial (kr) som slumpmässiga. Majas utgifter har väntevärdet 545 och standardavvikelsen 60, Joels väntevärdet 495 och standardavvikelsen 55. Utgifterna har korrelationen 0,65 och kan beskrivas med en tvådimensionell normalfördelning.

- a) Vad är sannolikheten att deras sammanlagda utgifter under en månad överstiger 1100 kr?
- b) Joel lägger 485 kr under en viss månad. Vad är sannolikheten att Majas utgifter samma månad överstiger 600 kr?

Lösning

Modell. ξ_1 = Majas utgifter, ξ_2 = Joels; (ξ_1, ξ_2) tvådimensionellt normalfördelat med $\mu_1 = 545$, $\sigma_1 = 60$, $\mu_2 = 495$, $\sigma_2 = 55$ och $\rho = 0,65$.

a) Summan är normalfördelat (kompendiets sats 1.21), med variansen från **variansen av en summa**:

$$D(\xi_1 + \xi_2) = \sqrt{60^2 + 55^2 + 2 \cdot 0,65 \cdot 60 \cdot 55} = \sqrt{10915} \approx 104,5$$

$$P(\xi_1 + \xi_2 > 1100) = 1 - \Phi\left(\frac{1100 - 1040}{104,5}\right) = 1 - \Phi(0,57) \approx \boxed{0,284}$$

Utan avrundning av z blir det 0,283.

b) Kompendiets sats 1.23 med $x_2 = 485$ (rollerna ombytta, vi betingar på Joel):

$$\mu' = 545 + 0,65 \cdot \frac{60}{55} (485 - 495) \approx 537,9, \quad \sigma' = 60 \sqrt{1 - 0,65^2} \approx 45,6$$

$$P(\xi_1 > 600 \mid \xi_2 = 485) = 1 - \Phi\left(\frac{600 - 537,9}{45,6}\right) = 1 - \Phi(1,36) \approx \boxed{0,087}$$

Joel ligger under sitt väntevärde, och med positiv korrelation drar det ner förväntningen på Maja. Spridningen krymper också, från 60 till 45,6.

På räknaren. a) `normCdf(1100, ∞, 1040, √(10915))`, b) `normCdf(600, ∞, 537.9, 45.6)`.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

≡ Kompendiet, övning 1.19 – Bevis av den betingade fördelningen

Bevisa kompendiets sats 1.23: om (ξ_1, ξ_2) är tvådimensionellt normalfördelat är den betingade fördelningen för ξ_2 givet $\xi_1 = x_1$ en $N(\mu'_2, \sigma'_2)$ -fördelning med $\mu'_2 = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x_1 - \mu_1)$ och $\sigma'_2 = \sigma_2 \sqrt{1 - \rho^2}$.

Lösning

Standardisera. Skriv $u = (x_1 - \mu_1)/\sigma_1$ och $v = (x_2 - \mu_2)/\sigma_2$. Hakparentesen i frekvensfunktionen är då $u^2 + v^2 - 2\rho uv$.

Kvadratkomplettera i v.

$$u^2 + v^2 - 2\rho uv = (v - \rho u)^2 + (1 - \rho^2) u^2$$

Frekvensfunktionen delar sig därför i två faktorer:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-u^2/2} e^{-\frac{(v-\rho u)^2}{2(1-\rho^2)}}$$

Dela med marginalen. $\xi_1 \in N(\mu_1, \sigma_1)$, så $f_{\xi_1}(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-u^2/2}$. Faktorn $e^{-u^2/2}$ och σ_1 försvinner:

$$f_{\xi_2 | \xi_1 = x_1}(x_2) = \frac{f(x_1, x_2)}{f_{\xi_1}(x_1)} = \frac{1}{\sqrt{2\pi} \sigma_2 \sqrt{1-\rho^2}} e^{-\frac{(v-\rho u)^2}{2(1-\rho^2)}}$$

Känn igen fördelningen. Multiplicera in σ_2 i täljaren och nämnaren i exponenten:

$$\frac{(v-\rho u)^2}{1-\rho^2} = \frac{(x_2 - \mu_2 - \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1))^2}{\sigma_2^2(1-\rho^2)} = \frac{(x_2 - \mu'_2)^2}{(\sigma'_2)^2}$$

Det är frekvensfunktionen för $N(\mu'_2, \sigma'_2)$ med μ'_2 och σ'_2 som i satsen. $\square$

Resonemangsuppgift – det finns inget facit att jämföra med.

≡ Kompendiet, övning 1.20 – Bygga en tvådimensionell normalfördelning

Antag att ζ_1 och ζ_2 är oberoende och $N(0, 1)$. Finn konstanter $a_{11}, a_{12}, a_{21}, a_{22}$ så att $(a_{11}\zeta_1 + a_{12}\zeta_2, a_{21}\zeta_1 + a_{22}\zeta_2)$, flyttad till rätt väntevärden, har en tvådimensionell normalfördelning med $\mu_1 = 1, \mu_2 = 3, \sigma_1 = 3, \sigma_2 = 4$ och $\rho = 0,25$.

Lösning

Vad som krävs. Linjärkombinationerna har väntevärdet 0, så väntevärdena 1 och 3 måste läggas till som konstanter. Vi söker

$$V(\xi_1) = a_{11}^2 + a_{12}^2 = 9, \quad V(\xi_2) = a_{21}^2 + a_{22}^2 = 16, \quad \text{Cov}(\xi_1, \xi_2) = a_{11}a_{21} + a_{12}a_{22} = 0,2$$

Ett val. Låt ξ_1 bara bero av ζ_1 : $a_{11} = 3, a_{12} = 0$. Kovariansvillkoret ger $3a_{21} = 3$, alltså $a_{21} = 1$, och variansvillkoret $a_{22} = \sqrt{15}$:

$$\xi_1 = 1 + 3\zeta_1, \quad \xi_2 = 3 + \zeta_1 + \sqrt{15}\zeta_2$$

Kontroll: $V(\xi_2) = 1 + 15 = 16$ och $\rho = 3/(3 \cdot 4) = 0,25$.

Varför det blir en tvådimensionell normalfördelning. Varje linjärkombination $c_1\xi_1 + c_2\xi_2$ är en linjärkombination av de oberoende normalfördelade ζ_1 och ζ_2 och alltså normalfördelad. Enligt omvändningen i kompendiets sats 1.21 är (ξ_1, ξ_2) då tvådimensionellt normalfördelad.

Valet är inte unikt: alla a_{ij} som uppfyller de tre villkoren fungerar.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

≡ Kompendiet, övning 1.21 – Normalfördelade marginaler räcker inte

Låt $\xi_1 \in N(0, 1)$ och $\xi_2 = \zeta \xi_1$, där ξ_1 och ζ är oberoende och $P(\zeta = -1) = P(\zeta = 1) = 0,5$.

- Bestäm fördelningen för ξ_2 .
- Har (ξ_1, ξ_2) en tvådimensionell normalfördelning?
- Har (ξ_1, ξ_2) en kontinuerlig tvådimensionell fördelning?

Lösning

a) Betinga på ζ med **lagen om total sannolikhet** och använd att $-\xi_1 \in N(0, 1)$ av symmetri:

$$P(\xi_2 \leq x) = 0,5 P(\xi_1 \leq x) + 0,5 P(-\xi_1 \leq x) = 0,5 \Phi(x) + 0,5 \Phi(x) = \Phi(x)$$

Alltså $\xi_2 \in N(0, 1)$.

b) Nej. Med $\zeta = -1$ är $\xi_1 + \xi_2 = 0$, så $P(\xi_1 + \xi_2 = 0) = 0,5$. En normalfördelad variabel antar aldrig ett enskilt värde med positiv sannolikhet, så linjärkombinationen $\xi_1 + \xi_2$ är inte normalfördelad. Enligt kompendiets sats 1.21 är då (ξ_1, ξ_2)

inte tvådimensionellt normalfördelad – trots att båda marginalerna är $N(0, 1)$.

c) Nej. Punkten (ξ_1, ξ_2) ligger alltid på någon av linjerna $x_2 = x_1$ och $x_2 = -x_1$. De har arean noll, och en dubbelintegral över ett område med arean noll är noll. Ingen frekvensfunktion kan därför ge linjerna sannolikheten 1, så **fördelningen är inte kontinuerlig**.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

Se även

- [Normalfördelning](#)
- [Korrelation](#)
- [Kovarians](#)
- [Betingad fördelning och betingat väntevärde](#)
- [Stickprov i par](#)

Resurser

- [Wikipedia: Multivariate normal distribution](#) — avsnittet om bivariata fallet
-
-