

Korrelation

19 okt. 2026, 6 min läsning

UTKAST

#matematik #statistik #storhet

Kurs: S0008M Förkunskaper: Kovarians

Kovariansen mäter hur två stokastiska variabler samvarierar. Korrelationskoefficienten är den normerade varianten, som alltid ligger mellan -1 och 1 och mäter det linjära beroendet — korrelation noll betyder att linjärt samband saknas, men inte nödvändigtvis att variablerna är oberoende.

1. Definitionen

Definition 1.18 (kompendiet). **Korrelationskoefficienten** mellan ξ_1 och ξ_2 är

$$\rho(\xi_1, \xi_2) = \frac{\text{Cov}(\xi_1, \xi_2)}{\sqrt{V(\xi_1) V(\xi_2)}} = \frac{\text{Cov}(\xi_1, \xi_2)}{\sigma_1 \sigma_2}$$

Vi skriver ibland ρ_{12} , eller bara ρ när det är tydligt vilka variabler som avses. Divisionen med standardavvikelserna gör måttet dimensionslöst: det påverkas inte av om vi mäter i meter eller centimeter.

2. Gränserna

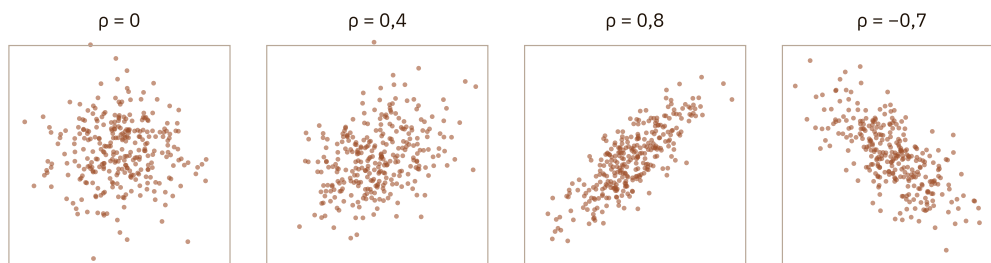
Det gäller alltid att

$$-1 \leq \rho(\xi_1, \xi_2) \leq 1$$

med likhet **om och endast om** ξ_2 är en linjär funktion av ξ_1 , alltså $\xi_2 = a\xi_1 + b$. Tecknet på ρ är tecknet på a .

ρ	Tolkning
nära 1	starkt positivt linjärt samband

ρ	Tolkning
nära 0	inget linjärt samband — variablerna är okorrelerade
nära -1	starkt negativt linjärt samband



Molnen visar observationspar från fyra fördelningar. Ju närmare $|\rho|$ är 1, desto tätare ligger punkterna kring en linje.

3. Okorrelerad är inte oberoende

⚠ $\rho = 0$ betyder inte oberoende

Oberoende variabler har alltid $\text{Cov} = 0$ och är alltså okorrelerade. Men det omvända gäller inte: ρ fångar bara **linjärt** beroende. Är (ξ, η) likformigt fördelad på en cirkelskiva är $\rho = 0$ av symmetriskäl, men variablerna är beroende — vet vi att ξ är nära 1 måste η vara nära 0.

Undantaget är den **tvådimensionella normalfördelningen**, där okorrelerade variabler också är oberoende.

4. Hur ρ påverkas av linjära omskrivningar

Ur räknereglererna för kovarians följer, för konstanter $a > 0$ och b :

$$\rho(a\xi_1 + b, \xi_2) = \rho(\xi_1, \xi_2), \quad \rho(-\xi_1, \xi_2) = -\rho(\xi_1, \xi_2)$$

Att skala om eller flytta en variabel ändrar inte styrkan i sambandet, men att byta tecken vänder riktningen.

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 1 – Kovarians och korrelation för omskrivna variabler

Slumpvariablerna ξ_1 och ξ_2 har väntevärden μ_1, μ_2 och standardavvikelser σ_1, σ_2 . Låt $\rho = \rho(\xi_1, \xi_2)$. Använd definitionerna för att bestämma kovarians och korrelation nedan, och kommentera om korrelationen beter sig som ett mått på linjärt samband.

- $\text{Cov}(\xi_1, \xi_1)$ och $\rho(\xi_1, \xi_1)$
- $\text{Cov}(\xi_1, 2\xi_1)$ och $\rho(\xi_1, 2\xi_1)$
- $\text{Cov}(\xi_1 + a, \xi_2)$ och $\rho(\xi_1 + a, \xi_2)$, där a är en konstant
- $\text{Cov}(\xi_1, -\xi_2)$ och $\rho(\xi_1, -\xi_2)$
- $\text{Cov}(\xi_1 + \xi_2, \xi_3)$ och $\rho(\xi_1 + \xi_2, \xi_3)$, där ξ_3 har väntevärde μ_3 och standardavvikelse σ_3 , och $\rho(\xi_1, \xi_3) = \rho_{13}, \rho(\xi_2, \xi_3) = \rho_{23}$

Lösning

a) $\text{Cov}(\xi_1, \xi_1) = E((\xi_1 - \mu_1)^2) = \boxed{\sigma_1^2}$, och $\rho = \sigma_1^2/(\sigma_1\sigma_1) = \boxed{1}$. En variabel har perfekt linjärt samband med sig själv.

b) $2\xi_1$ har väntevärde $2\mu_1$ och standardavvikelse $2\sigma_1$:

$$\text{Cov}(\xi_1, 2\xi_1) = E((\xi_1 - \mu_1) \cdot 2(\xi_1 - \mu_1)) = \boxed{2\sigma_1^2}, \quad \rho = \frac{2\sigma_1^2}{\sigma_1 \cdot 2\sigma_1} = \boxed{1}$$

Kovariansen ändras med skalan, korrelationen inte.

c) $\xi_1 + a$ har väntevärde $\mu_1 + a$, så avvikelsen är oförändrad, $(\xi_1 + a) - (\mu_1 + a) = \xi_1 - \mu_1$:

$$\text{Cov}(\xi_1 + a, \xi_2) = \boxed{\rho \sigma_1 \sigma_2}, \quad \rho(\xi_1 + a, \xi_2) = \boxed{\rho}$$

d) Avvikelsen för $-\xi_2$ är $-(\xi_2 - \mu_2)$ och standardavvikelsen fortfarande σ_2 :

$$\text{Cov}(\xi_1, -\xi_2) = \boxed{-\rho \sigma_1 \sigma_2}, \quad \rho(\xi_1, -\xi_2) = \boxed{-\rho}$$

Sambandet är lika starkt men går åt andra hållet.

e) Kovariansen är additiv i första argumentet:

$$\text{Cov}(\xi_1 + \xi_2, \xi_3) = \text{Cov}(\xi_1, \xi_3) + \text{Cov}(\xi_2, \xi_3) = \boxed{\rho_{13} \sigma_1 \sigma_3 + \rho_{23} \sigma_2 \sigma_3}$$

Med **kompendiets sats 1.17** är $V(\xi_1 + \xi_2) = \sigma_1^2 + \sigma_2^2 + 2\rho \sigma_1 \sigma_2$, så

$$\rho(\xi_1 + \xi_2, \xi_3) = \boxed{\frac{\rho_{13} \sigma_1 \sigma_3 + \rho_{23} \sigma_2 \sigma_3}{\sigma_3 \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho \sigma_1 \sigma_2}}}$$

Kommentar. I a)–d) betar sig ρ som väntat för ett mått på linjärt samband: ± 1 vid exakt linjärt samband, oförändrad vid flyttning och skalning, och teckenbyte när den ena variabeln byter tecken.

Facit: a) σ_1^2 resp. 1; b) $2\sigma_1^2$ resp. 1; c) $\rho\sigma_1\sigma_2$ resp. ρ ; d) $-\rho\sigma_1\sigma_2$ resp. $-\rho$; e) som ovan.

≡ Kompendiet, övning 1.16 – Räkneregler för kovarians och korrelation

Verifiera följande egenskaper, där a och b är konstanter:

a) $\text{Cov}(a\xi_1 + b, \xi_2) = a \text{Cov}(\xi_1, \xi_2)$

b) $\text{Cov}(\xi_1 + \xi_2, \xi_3) = \text{Cov}(\xi_1, \xi_3) + \text{Cov}(\xi_2, \xi_3)$

c) $\rho(\xi_1, -\xi_2) = -\rho(\xi_1, \xi_2)$

d) $\rho(\xi_1, a\xi_1 + b)$ är lika med -1 eller 1 (för $a \neq 0$).

Lösning

Använd $\text{Cov}(\xi_1, \xi_2) = E(\xi_1 \xi_2) - \mu_1 \mu_2$ och att väntevärdet är linjärt.

a) $E((a\xi_1 + b)\xi_2) = aE(\xi_1\xi_2) + b\mu_2$ och $E(a\xi_1 + b)E(\xi_2) = a\mu_1\mu_2 + b\mu_2$. Differensen är $a(E(\xi_1\xi_2) - \mu_1\mu_2) = a\text{Cov}(\xi_1, \xi_2)$.

b) $E((\xi_1 + \xi_2)\xi_3) - (\mu_1 + \mu_2)\mu_3 = (E(\xi_1\xi_3) - \mu_1\mu_3) + (E(\xi_2\xi_3) - \mu_2\mu_3)$, vilket är högerledet.

c) Med a) och $a = -1$, $b = 0$ (rollerna ombytta, kovariansen är symmetrisk) är $\text{Cov}(\xi_1, -\xi_2) = -\text{Cov}(\xi_1, \xi_2)$. Standardavvikelsen för $-\xi_2$ är σ_2 , så nämnaren i ρ är oförändrad och tecknet byts.

d) Med a) är $\text{Cov}(\xi_1, a\xi_1 + b) = aV(\xi_1) = a\sigma_1^2$, och $D(a\xi_1 + b) = |a|\sigma_1$:

$$\rho(\xi_1, a\xi_1 + b) = \frac{a\sigma_1^2}{\sigma_1 \cdot |a|\sigma_1} = \frac{a}{|a|} = \begin{cases} 1, & a > 0 \\ -1, & a < 0 \end{cases}$$

Ett exakt linjärt samband ger alltså alltid $|\rho| = 1$, med tecknet efter lutningen.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

≡ Kompendiet, övning 1.17 – Okorrelerade men beroende

Låt (ξ_1, ξ_2) vara likformigt fördelad på cirkelskivan med centrum i origo och radien 1. Bestäm $\text{Cov}(\xi_1, \xi_2)$. Är ξ_1 och ξ_2 okorrelerade? Är de oberoende?

Lösning

Kovariansen. Arealen är π , så $f(x_1, x_2) = 1/\pi$ på skivan. Skivan är symmetrisk kring båda axlarna, och integranderna x_1 och x_1x_2 är udda i x_1 , så integralerna över skivan blir noll:

$$\mu_1 = \mu_2 = 0, \quad E(\xi_1\xi_2) = \iint_{x_1^2+x_2^2 \leq 1} \frac{x_1x_2}{\pi} dx_1 dx_2 = 0$$

Alltså $\boxed{\text{Cov}(\xi_1, \xi_2) = 0}$, och variablerna är **okorrelerade**.

Oberoende? Nej. Marginalerna är $f_{\xi_1}(x_1) = \frac{2}{\pi} \sqrt{1 - x_1^2}$ på $[-1, 1]$, och likadant för ξ_2 (se [Marginalfördelning](#)). I punkten $(0,9; 0,9)$, som ligger utanför skivan, är $f = 0$ men produkten av marginalerna positiv. Alltså är $f(x_1, x_2) \neq f_{\xi_1}(x_1)f_{\xi_2}(x_2)$, och variablerna är **beroende**.

Det här är exemplet i avsnitt 3: vet vi att ξ_1 är nära 1 måste ξ_2 vara nära 0. Beroendet är inte linjärt, och därför syns det inte i ρ .

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

Se även

- [Kovarians](#)
- [Tvådimensionell normalfördelning](#)
- [Flerdimensionella fördelningar](#)
- [Oberoende stokastiska variabler](#)

Resurser

- [Khan Academy: Correlation coefficient](#) 
- [Wikipedia: Korrelation](#) 