

Kovarians

19 okt. 2026, 5 min läsning

UTKAST

#matematik #sannolikhetslära #storhet

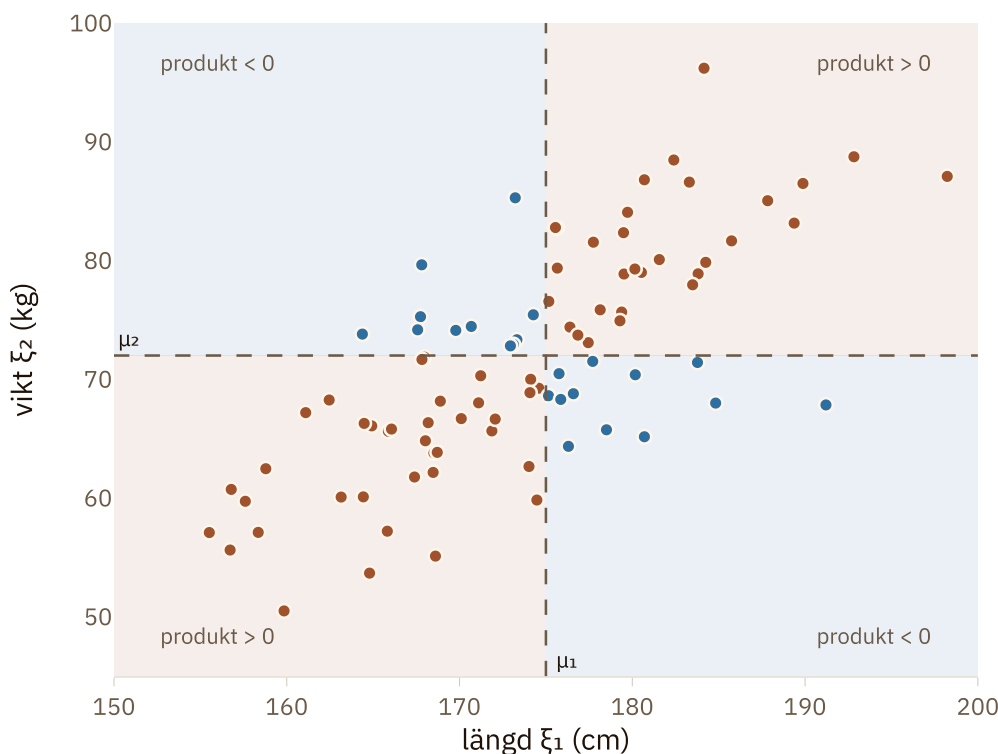
Kurs: S0008M Förkunskaper: Flerdimensionella fördelningar, Varians och standardavvikelse

Kovariansen mäter hur två stokastiska variabler samvarierar: den är positiv om stora värden på den ena tenderar att åtföljas av stora värden på den andra, och negativ om de tenderar att gå åt olika håll. Den behövs för att räkna ut variansen av en summa av beroende variabler.

1. Idén

Låt ξ_1 och ξ_2 vara längd och vikt för en slumpvald person, med väntevärden μ_1 och μ_2 . En person över medellängd väger oftast över medelvikten, och en person under medellängd oftast under. Avvikelseerna $\xi_1 - \mu_1$ och $\xi_2 - \mu_2$ har alltså oftast samma tecken, så produkten $(\xi_1 - \mu_1)(\xi_2 - \mu_2)$ blir i genomsnitt positiv.

Vid ett negativt samband har avvikelseerna oftast olika tecken och produkten blir i genomsnitt negativ. Saknas linjärt samband tar positiva och negativa produkter ut varandra.



Punkter i de rödbruna kvadranterna ger en positiv produkt $(\xi_1 - \mu_1)(\xi_2 - \mu_2)$, punkter i de blå en negativ. Här dominerar de rödbruna, så kovariansen är positiv. Håll musen över en punkt för att se produkten.

2. Definitionen

Definition 1.16 (kompendiet). **Kovariansen** mellan ξ_1 och ξ_2 är

$$\text{Cov}(\xi_1, \xi_2) = E((\xi_1 - \mu_1)(\xi_2 - \mu_2)), \quad \mu_i = E(\xi_i)$$

För räkning är omskrivningen

$$\text{Cov}(\xi_1, \xi_2) = E(\xi_1 \xi_2) - \mu_1 \mu_2$$

oftast smidigast, där $E(\xi_1 \xi_2)$ räknas som en dubbelsumma eller dubbelintegral över den **simultana fördelningen**.

Två specialfall

- **Med sig själv:** $\text{Cov}(\xi, \xi) = E((\xi - \mu)^2) = V(\xi)$. Variansen är kovariansen av en variabel med sig själv.
- **Oberoende:** om ξ_1 och ξ_2 är oberoende är $E(\xi_1 \xi_2) = \mu_1 \mu_2$, så $\text{Cov}(\xi_1, \xi_2) = 0$.

3. Räkneregler

För konstanter a och b :

$$\begin{aligned}\text{Cov}(\xi_1, \xi_2) &= \text{Cov}(\xi_2, \xi_1) \\ \text{Cov}(a\xi_1 + b, \xi_2) &= a \text{Cov}(\xi_1, \xi_2) \\ \text{Cov}(\xi_1 + \xi_2, \xi_3) &= \text{Cov}(\xi_1, \xi_3) + \text{Cov}(\xi_2, \xi_3)\end{aligned}$$

En additiv konstant flyttar variabeln men inte avvikelserna från väntevärdet, och därför försvinner b .

4. Variansen av en summa

Sats 1.17 (kompendiet).

$$V(\xi_1 + \xi_2) = V(\xi_1) + V(\xi_2) + 2 \text{Cov}(\xi_1, \xi_2)$$

och, eftersom $V(-\xi_2) = V(\xi_2)$ och $\text{Cov}(\xi_1, -\xi_2) = -\text{Cov}(\xi_1, \xi_2)$,

$$V(\xi_1 - \xi_2) = V(\xi_1) + V(\xi_2) - 2 \text{Cov}(\xi_1, \xi_2)$$

Varför. Utveckla kvadraten i definitionen av variansen:

$$V(\xi_1 + \xi_2) = E(((\xi_1 - \mu_1) + (\xi_2 - \mu_2))^2) = E((\xi_1 - \mu_1)^2) + E((\xi_2 - \mu_2)^2) + 2E((\xi_1 - \mu_1)(\xi_2 - \mu_2))$$

Är variablerna oberoende försvinner kovarianstermen och vi får tillbaka regeln att varianser adderas, se [Linjära funktioner av stokastiska variabler](#).

Positivt beroende ökar spridningen i summan

Är kovariansen positiv drar variablerna åt samma håll och summan sprider mer än om de vore oberoende. För en differens är det tvärtom — det är därför **stickprov i par** ger snävare intervall: mätningarna i ett par är positivt korrelerade, och differensen tar bort den gemensamma variationen.

5. Skalberoende

Kovariansen har samma enhet som produkten av variablernas enheter. Mäter vi längd i meter i stället för centimeter blir kovariansen 100 gånger mindre, utan att sambandet har ändrats. För att få ett mått som inte beror på skalan dividerar vi med standardavvikelse, och det ger **korrelationskoefficienten**:

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

Kovariansmatris

För flera variabler samlas varianser och kovarianser i en **kovariansmatris** C med elementen $c_{ij} = \text{Cov}(\xi_i, \xi_j)$. I två dimensioner är

$$C = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

Vanliga fel

- **Cov = 0 betyder inte oberoende.** Oberoende ger kovariansen 0, men inte omvänt. Med $\xi \in R(-1, 1)$ och $\eta = \xi^2$ är $\text{Cov}(\xi, \eta) = E(\xi^3) - 0 \cdot E(\xi^2) = 0$, fast η bestäms helt av ξ .
- **Glömd faktor 2.** $V(\xi_1 + \xi_2)$ har termen 2Cov , och för $a\xi_1 + b\xi_2$ blir den $2ab \text{Cov}(\xi_1, \xi_2)$.
- **Tecknet vid differens.** I $V(\xi_1 - \xi_2)$ byter bara kovarianstermen tecken – varianserna adderas fortfarande.
- $E(\xi_1 \xi_2) = \mu_1 \mu_2$ **utan oberoende.** Produktregeln gäller bara för oberoende variabler; annars räknas $E(\xi_1 \xi_2)$ som dubbelsumma eller dubbelintegral.

6. 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 3 – Hur stor kan variansen av en summa bli?

Låt ξ_1 och ξ_2 ha varianserna 9 respektive 25. Om de är beroende kan variansen av $\xi_1 + \xi_2$ vara antingen större eller mindre än 34, men inte hur stor eller liten som helst. Hur stor kan den högst vara? Hur liten kan den minst vara?

Lösning

Kompendiets sats 1.17 med $\text{Cov}(\xi_1, \xi_2) = \rho \sigma_1 \sigma_2 = 3 \cdot 5 \cdot \rho = 15\rho$:

$$V(\xi_1 + \xi_2) = 9 + 25 + 2 \cdot 15\rho = 34 + 30\rho$$

Korrelationen ligger alltid i $-1 \leq \rho \leq 1$, så

$$4 \leq V(\xi_1 + \xi_2) \leq 64$$

Maxvärdet $64 = (3 + 5)^2$ nås när $\rho = 1$ och variablerna drar exakt åt samma håll; minvärdet $4 = (5 - 3)^2$ när $\rho = -1$ och de delvis tar ut varandra.

Facit: $4 \leq V(\xi_1 + \xi_2) \leq 64$.

≡ Kompendiet, övning 1.15 – Samma fråga med varianserna 4 och 9

Låt ξ_1 och ξ_2 ha varianserna 4 respektive 9. Hur stor kan variansen av $\xi_1 + \xi_2$ högst vara, och hur liten kan den minst vara?

Lösning

Nu är $\sigma_1 = 2$ och $\sigma_2 = 3$, så $\text{Cov}(\xi_1, \xi_2) = 6\rho$ och

$$V(\xi_1 + \xi_2) = 4 + 9 + 12\rho = 13 + 12\rho$$

Med $-1 \leq \rho \leq 1$:

$$1 \leq V(\xi_1 + \xi_2) \leq 25$$

Gränserna är $(\sigma_2 - \sigma_1)^2$ och $(\sigma_1 + \sigma_2)^2$, precis som i lektionsuppgiften ovan.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

Se även

- Korrelation
- Flerdimensionella fördelningar
- Varians och standardavvikelse
- Linjära funktioner av stokastiska variabler
- Tvådimensionell normalfördelning

Resurser

- [Khan Academy: Covariance and the regression line](#)
 - [Wikipedia: Kovarians](#)
-
-