

Stickprov i par

19 okt. 2026, 6 min läsning

UTKAST

#matematik #statistik #metod

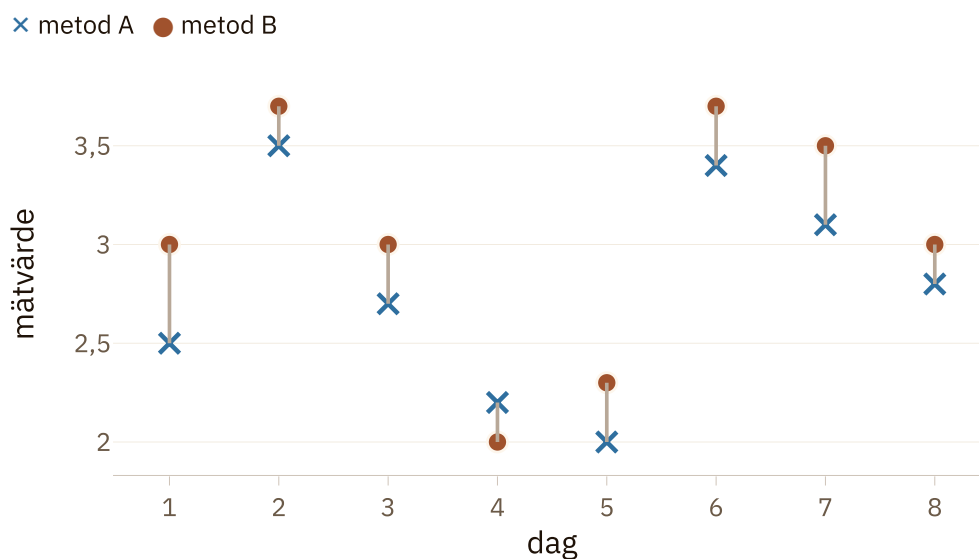
Kurs: S0008M Förkunskaper: Konfidenstervall, t-fördelningen

Stickprov i par används när observationerna kommer parvis, till exempel två mätningar på samma individ. Vi bildar differensen inom varje par och analyserar dessa differenser som ett enda stickprov, vilket räknar bort variationen mellan individer.

1. Situationen

Vi vill jämföra två metoder, A och B, och planerar försöket så att båda används på samma objekt: samma dag, samma patient, samma vajer delad i två bitar. Då får vi par av mätvärden (x_i, y_i) , $i = 1, \dots, n$.

Objekten kan skilja sig mycket från varandra — en dag ger höga värden, en annan låga — men det intressanta är skillnaden *inom* varje par.



I figuren varierar nivån kraftigt från dag till dag, men B ligger nästan alltid lite över A. Det är den systematiska skillnaden vi vill skatta.

2. Modellen

x_i observation från $N(\mu_i, \sigma_1)$, y_i observation från $N(\mu_i + \Delta, \sigma_2)$, $i = 1, \dots, n$

- μ_i är nivå för objekt i . Den får variera fritt mellan objekten, och vi behöver inte skatta den.
- Δ är den **förväntade skillnaden** mellan metoderna — samma för alla objekt. Det är parametern vi vill veta något om.

Bilda differenserna $z_i = y_i - x_i$. Nivån μ_i tar ut sig:

z_i är observationer från $N(\Delta, \sigma)$, $i = 1, \dots, n$, oberoende

Problemet har blivit ett enda stickprov med okänt väntevärde Δ och okänd standardavvikelse — precis situationen i **t-fördelningen**.

3. Konfidensintervallet

Med $\bar{z}$ och s_z beräknade ur differenserna:

$$I_{\Delta} = \left[\bar{z} - t_{\alpha/2}(n-1) \frac{s_z}{\sqrt{n}}, \bar{z} + t_{\alpha/2}(n-1) \frac{s_z}{\sqrt{n}} \right]$$

Tolka med nollan

- Ingår **inte** 0 i I_{Δ} : vi kan med den givna säkerheten påstå att metoderna skiljer sig i genomsnitt, och intervallet säger hur mycket.
- Ingår 0: vi kan inte påvisa någon skillnad. Det betyder inte att skillnaden är noll — bara att datamaterialet inte räcker för att visa den.

På räknaren: lägg mätvärdena i två kolumner i Lists & Spreadsheet, bilda differenserna i en tredje med kolumnformeln `=b[] - a[]`, och kör t Interval med **Data** på den kolumnen.

Varför paren lönar sig

Variationen mellan objekten, $\mu_1, \dots, \mu_n$, försvinner helt ur differenserna. Hade vi i stället mätt metod A på några objekt och metod B på andra hade den variationen hamnat i spridningen, och intervallet för skillnaden blivit mycket bredare. Jämför **Två stickprov**.

⚠ Vanliga fel

- **Analysera paren som två stickprov.** Då hamnar variationen mellan objekten i s och intervallet blir onödigt brett.
- **Blanda ordningen i differensen.** Bestäm $z_i = x_i - y_i$ en gång och tolka tecknet därefter.

4. Lösta lektionsuppgifter

≡ Lektion 12, övning 5 – Två mätmetoder under åtta dagar

Två mätmetoder, A och B, används samma dag under åtta dagar:

Dag	1	2	3	4	5	6
Metod A	2,5	3,5	2,7	2,2	2,0	3,4
Metod B	3,0	3,7	3,0	2,0	2,3	3,5

- Pricka in mätvärdena dag för dag i ett diagram. Tror du att metoderna skiljer sig åt?
- Anta att mätvärdet för metod A dag i är en observation från $N(\mu_i, \sigma)$ och för metod B från $N(\mu_i + \Delta, \sigma)$. Tolk Δ och $\mu_1, \dots, \mu_8$.
- Kan vi med 95 % säkerhet påstå att metoderna skiljer sig i genomsnitt? Hur stor är i så fall skillnaden?

✎ Lösning

- Se figuren i avsnitt 1. Nivån varierar mycket mellan dagarna, men B ligger över A alla dagar utom dag 4. Det tyder på en systematisk skillnad.
- μ_i är den förväntade nivån dag i — det som gör att vissa dagar ger höga värden och andra låga. Det är rimligt att de skiljer sig, och figuren visar att de gör det. Δ är den förväntade skillnaden mellan metod B och metod A, densamma alla dagar.
- Differenserna $z_i = y_i - x_i$ (B – A):

0,5 0,2 0,3 - 0,2 0,3 0,3 0,4 0,2

är observationer från $N(\Delta, \sigma_z)$. $\bar{z} = 0,25$, $s_z \approx 0,207$, $t_{0,025}(7) = 2,365$:

$$I_\Delta = [0,25 \mp 2,365 \cdot \frac{0,207}{\sqrt{8}}] \approx [0,08; 0,42]$$

0 ingår inte. Med 95 % säkerhet finns en genomsnittlig skillnad: metod B ger i genomsnitt mellan 0,08 och 0,42 högre mätvärden än metod A.

Rimlighetskontroll. Differenserna är exakt observationerna i [lektion 12, övning 2](#), så intervallet måste bli detsamma.

På räknaren. t Interval med **Data** på differenskolumnen, **C Level** 0.95 .

Facit: [0,07; 0,43], avrundat utåt; metod B ger i genomsnitt mellan 0,07 och 0,43 högre mätvärden än metod A.

≡ Lektion 13, övning 1 – Torsionsprov på stålvaror

För att studera om två maskiner, A och B, för torsionsprov på stålvaror ger olika resultat valdes sex varor av olika typ ut och delades i två likadana bitar. Slumpen avgjorde vilken bit som testades i vilken maskin.

Typ av varor	1	2	3	4	5	6
Maskin A	27	35	38	28	40	42
Maskin B	25	34	39	26	37	42

Kan vi påstå att maskinerna i genomsnitt ger olika resultat? Hur stor är i så fall skillnaden? Beräkna och tolka ett 90 % konfidensintervall, och redovisa modellantagandena.

 Lösning

Modell. Mätvärdet för vajer i i maskin A är en observation från $N(\mu_i + \Delta, \sigma_1)$ och i maskin B från $N(\mu_i, \sigma_2)$, alla oberoende. μ_i är vajertypens nivå och Δ den förväntade skillnaden A – B. Stickprov i par, eftersom varje vajer testas i båda maskinerna.

Differenser $z_i = A - B$:

2 1 - 1 2 3 0

observationer från $N(\Delta, \sigma)$. $\bar{z} \approx 1,167$, $s_z \approx 1,472$, $t_{0,05}(5) = 2,015$:

$$I_{\Delta} = [1,167 \mp 2,015 \cdot \frac{1,472}{\sqrt{6}}] = [1,167 \mp 1,211] \approx [-0,04; 2,38]$$

Tolkning. 0 ingår i intervallet, så vi kan inte med 90 % säkerhet påvisa någon skillnad i förväntat mätresultat mellan maskinerna. Den nedre gränsen ligger nära 0, så fler observationer skulle behövas för ett säkrare besked.

På räknaren. t Interval på differenserna med C Level 0.9.

Facit: $[-0,05; 2,38]$, avrundat utåt; ingen skillnad kan påvisas med 90 % säkerhet, fler observationer rekommenderas.

☰ Lektion 13, övning 3b (Vännman 8.23 b) – Blodtryck före och efter

I ett försök mäts blodtrycket hos 25 patienter före (x) och efter (y) behandling med en medicin. Resultat: $\bar{x} = 149$, $s_x = 8,1$, $\bar{y} = 150,9$, $s_y = 9,5$, och för differenserna $z_i = y_i - x_i$: $\bar{z} = 1,9$, $s_z = 1,6$. Beräkna ett 95 % konfidensintervall för den förväntade förändringen och tolka det.

 Lösning

Modell. x_i observation från $N(\mu_i, \sigma_1)$, y_i från $N(\mu_i + \Delta, \sigma_2)$; differenserna z_i observationer från $N(\Delta, \sigma)$, $i = 1, \dots, 25$. Stickprov i par – samma patient mäts två gånger.

$$t_{0,025}(24) = 2,064:$$

$$I_{\Delta} = [1,9 \mp 2,064 \cdot \frac{1,6}{\sqrt{25}}] = [1,9 \mp 0,66] \approx [1,24; 2,56]$$

Tolkning. 0 ingår inte. Med 95 % säkerhet höjer medicinen blodtrycket i genomsnitt med mellan 1,24 och 2,56 enheter.

Observera att $s_z = 1,6$ är mycket mindre än s_x och s_y : variationen mellan patienter är stor, men förändringen hos varje patient är liten och stabil. Deluppgift a), samma fråga med två stickprov, finns i **Två stickprov**.

Uppgiften (Vännman 8.23 b, i 2:a uppl. 8.25 b) är återgiven ur det handskrivna lösningsförslaget, och intervallet stämmer med det.

Se även

- **Två stickprov**
- **t-fördelningen**
- **Konfidensintervall**
- **Hypotesprövning**

Resurser

- [Wikipedia: Paired difference test](#) 