

Beskrivande statistik

Descriptive Statistics

Mathias Johansson Kristofer Söderström

2025-12-15

Innehållsförteckning

1	Vad är beskrivande statistik?	2
2	Centralmått	2
2.1	Medelvärde (Mean)	2
3	Median	2
4	Typvärde (Mode)	2
5	När använda vilket centralmått?	2
6	Medelvärde vs Median	3
7	Spridningsmått	3
7.1	Variationsbredd (Range)	3
8	Varians (Variance)	3
9	Standardavvikelse (Standard deviation)	3
10	Kvartiler (Quartiles)	4
11	Percentiler och kvantiler	4
12	Beräkna kvartiler	4
13	Boxplot (lådagram)	4
14	Skevhet (Skewness)	4
15	Sammanfattning - fem siffror	5
16	Praktisk tillämpning	5
17	Referenser	5

1 Vad är beskrivande statistik?

Metoder för att sammanfatta och beskriva data

- **Centralmått:** Var ligger "mitten"?
- **Spridningsmått:** Hur mycket varierar data?
- **Fördelning:** Hur ser data ut?

2 Centralmått

2.1 Medelvärde (Mean)

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Exempel: 2, 4, 4, 5, 10

$$\bar{x} = \frac{2 + 4 + 4 + 5 + 10}{5} = \frac{25}{5} = 5$$

3 Median

Mittvärdet när data är sorterat

Udda antal: Mittobservationen **Jämnt antal:** Medelvärdet av de två mittvärdarna

Exempel: 2, 4, 4, 5, 10 Median = 4 (tredje värdet)

4 Typvärde (Mode)

Det värde som förekommer oftast

Exempel: 2, 4, 4, 5, 10 Typvärde = 4 (förekommer 2 gånger)

Kan ha flera typvärden (multimodal fördelning)

5 När använda vilket centralmått?

Datatyp	Rekommenderat mått
Nominal	Typvärde
Ordinal	Median
Intervall/Kvot	Medelvärde*

*Om fördelningen är symmetrisk och utan extremvärden

6 Medelvärde vs Median

Skev fördelning eller extremvärden? → Använd median

Exempel: Inkomster - 5 personer: 25k, 28k, 30k, 32k, 200k - Medelvärde: 63k - Median: 30k

Medianen ger en mer "typisk" bild!

7 Spridningsmått

7.1 Variationsbredd (Range)

$$\text{Range} = x_{\max} - x_{\min}$$

Exempel: 2, 4, 4, 5, 10 Range = 10 - 2 = 8

Känslig för extremvärden

8 Varians (Variance)

Genomsnittlig kvadrerad avvikelse från medelvärdet

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Exempel: 2, 4, 4, 5, 10 (medel = 5)

$$s^2 = \frac{(2-5)^2 + (4-5)^2 + (4-5)^2 + (5-5)^2 + (10-5)^2}{4}$$

$$s^2 = \frac{9+1+1+0+25}{4} = 9$$

9 Standardavvikelse (Standard deviation)

Roten ur variansen - samma enhet som data

$$s = \sqrt{s^2}$$

Exempel: $s = \sqrt{9} = 3$

Tumregel: Ca 68% av data ligger inom ± 1 std från medelvärdet

10 Kvartiler (Quartiles)

Delar data i fyra lika delar

- **Q1** (25:e percentilen): 25% under
- **Q2** (median): 50% under
- **Q3** (75:e percentilen): 75% under

Interkvartilt intervall (IQR): $Q3 - Q1$

11 Percentiler och kvantiler

Percentil: Värdet under vilket en viss procent av data ligger

- 10:e percentilen: 10% under
- 90:e percentilen: 90% under

Kvantil: Generellt begrepp - Kvartiler = 4 delar - Deciler = 10 delar - Percentiler = 100 delar

12 Beräkna kvartiler

Data: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

- $Q1 = 4$ (25%)
- $Q2 = 7$ (50%, median)
- $Q3 = 10$ (75%)
- $IQR = 10 - 4 = 6$

13 Boxplot (lådagram)

Visuell sammanfattning av fördelningen

Q1 Q2 Q3

Min Max
 IQR

14 Skevhet (Skewness)

Mått på fördelningens asymmetri

- **Positivt skev:** Svans åt höger ($medel > median$)
- **Negativt skev:** Svans åt vänster ($medel < median$)
- **Symmetrisk:** Medel = Median

15 Sammanfattning - fem siffror

Five-number summary:

1. Minimum
2. Q1
3. Median (Q2)
4. Q3
5. Maximum

Ger en snabb överblick av datafördelningen!

16 Praktisk tillämpning

När du presenterar enkätresultat:

1. **Ange alltid n** (antal svar)
2. **Välj rätt centralmått** för datatypen
3. **Visa spridning** (std eller IQR)
4. **Visualisera** med lämpligt diagram

17 Referenser

- Descriptive statistics
- Mean
- Median
- Mode (statistics)
- Standard deviation
- Variance
- Quartile
- Box plot
- Skewness