

# Inferentiell statistik

## Z-värden, p-värden och felmarginaler

Mathias Johansson

Kristofer Söderström

2025-12-15

### Innehållsförteckning

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introduktion</b>                                       | <b>2</b> |
| <b>2</b> | <b>Z-värden (Z-scores)</b>                                | <b>2</b> |
| 2.1      | Vad är ett Z-värde? . . . . .                             | 2        |
| 2.2      | Tolkning . . . . .  | 2        |
| 2.3      | Z-värden och normalfördelningen . . . . .                 | 2        |
| 2.4      | Kritiska Z-värden för konfidensnivåer . . . . .           | 2        |
| 2.5      | Exempel: IQ-test . . . . .                                | 3        |
| <b>3</b> | <b>P-värden (P-values)</b>                                | <b>3</b> |
| 3.1      | Vad är ett p-värde? . . . . .                             | 3        |
| 3.2      | Tolkning . . . . .  | 3        |
| 3.3      | Vanliga missförstånd . . . . .                            | 3        |
| 3.4      | Signifikansnivå ( ) . . . . .                             | 3        |
| 3.5      | Exempel: Myntkast . . . . .                               | 3        |
| <b>4</b> | <b>Felmarginaler (Margin of Error)</b>                    | <b>4</b> |
| 4.1      | Vad är en felmarginal? . . . . .                          | 4        |
| 4.2      | Konfidensintervall . . . . .                              | 4        |
| 4.3      | Faktorer som påverkar felmarginalen . . . . .             | 4        |
| 4.4      | Exempel: Opinionsundersökning . . . . .                   | 4        |
| <b>5</b> | <b>Samband mellan begreppen</b>                           | <b>5</b> |
| 5.1      | Z-värden och p-värden . . . . .                           | 5        |
| 5.2      | Z-värden och felmarginaler . . . . .                      | 5        |
| <b>6</b> | <b>Praktisk tillämpning</b>                               | <b>5</b> |
| 6.1      | Hypotestest i 5 steg . . . . .                            | 5        |
| 6.2      | Exempel: Är studenter längre än genomsnittet? . . . . .   | 6        |
| <b>7</b> | <b>Vanliga misstag att undvika</b>                        | <b>6</b> |
| 7.1      | 1. P-hacking . . . . .                                    | 6        |
| 7.2      | 2. Förväxla statistisk och praktisk signifikans . . . . . | 6        |
| 7.3      | 3. Ignorera urvalsstorlek . . . . .                       | 6        |
| 7.4      | 4. Feltolkning av konfidensintervall . . . . .            | 6        |
| <b>8</b> | <b>Viktiga begrepp</b>                                    | <b>6</b> |

|                              |          |
|------------------------------|----------|
| <b>9 Interaktiva verktyg</b> | <b>7</b> |
| <b>10 Vidare läsning</b>     | <b>7</b> |
| <b>11 Referenser</b>         | <b>7</b> |

## 1 Introduktion

Inferentiell statistik handlar om att dra slutsatser om en population baserat på ett urval. Tre centrala begrepp är Z-värden, p-värden och felmarginaler.

## 2 Z-värden (Z-scores)

### 2.1 Vad är ett Z-värde?

Ett Z-värde anger hur många standardavvikelser ett värde ligger från medelvärdet i en normalfördelning.

$$z = \frac{x - \mu}{\sigma}$$

- $z$  = Z-värde
- $x$  = Observerat värde
- $\mu$  = Medelvärde
- $\sigma$  = Standardavvikelse

### 2.2 Tolkning

- **Z = 0:** Värdet är exakt på medelvärdet
- **Z = 1:** Värdet ligger 1 standardavvikelse över medelvärdet
- **Z = -1:** Värdet ligger 1 standardavvikelse under medelvärdet
- **Z = 2:** Värdet ligger 2 standardavvikelser över medelvärdet

### 2.3 Z-värden och normalfördelningen

I en normalfördelning: - **68%** av värdena ligger inom  $\pm 1$  standardavvikelse (Z mellan -1 och +1) - **95%** av värdena ligger inom  $\pm 1.96$  standardavvikelser (Z mellan -1.96 och +1.96) - **99%** av värdena ligger inom  $\pm 2.58$  standardavvikelser (Z mellan -2.58 och +2.58)

### 2.4 Kritiska Z-värden för konfidensnivåer

| Konfidensnivå | Z-värde     | Användning                  |
|---------------|-------------|-----------------------------|
| 90%           | 1.645       | Explorativa studier         |
| <b>95%</b>    | <b>1.96</b> | <b>Standard i forskning</b> |
| 99%           | 2.576       | Hög säkerhet                |
| 99.9%         | 3.291       | Mycket hög säkerhet         |

## 2.5 Exempel: IQ-test

IQ har medelvärde 100 och standardavvikelse 15.

Om en person har IQ 130:

$$z = \frac{130 - 100}{15} = \frac{30}{15} = 2$$

Detta betyder att personen ligger 2 standardavvikelser över medelvärdet, vilket är högre än ca 97.5% av populationen.

## 3 P-värden (P-values)

### 3.1 Vad är ett p-värde?

P-värdet är sannolikheten att få ett resultat som är lika extremt eller mer extremt än det observerade resultatet, *givet att nollhypotesen är sann*.

### 3.2 Tolkning

- $p < 0.001$ : Mycket starkt bevis mot nollhypotesen (\*\*\*)
- $p < 0.01$ : Starkt bevis mot nollhypotesen (\*\*)
- $p < 0.05$ : Måttligt bevis mot nollhypotesen (\*)
- $p = 0.05$ : Otillräckligt bevis för att förkasta nollhypotesen

### 3.3 Vanliga missförstånd

**FELAKTIGT:** “ $p = 0.05$  betyder att det är 5% chans att nollhypotesen är sann”

**KORREKT:** “ $p = 0.05$  betyder att om nollhypotesen vore sann, skulle vi se dessa data (eller mer extrema) i 5% av fallen”

### 3.4 Signifikansnivå ( )

Signifikansnivån (ofta  $= 0.05$ ) är tröskeln vi sätter *innan* studien för att bestämma om ett resultat ska betraktas som statistiskt signifikant.

- Om  $p < \alpha$  : Förkasta nollhypotesen (signifikant resultat)
- Om  $p \geq \alpha$  : Behåll nollhypotesen (icke-signifikant resultat)

### 3.5 Exempel: Myntkast

Nollhypotes: Myntet är rättvist ( $p = 0.5$  för krona)

Du kastar 100 gånger och får 65 kronor.

Fråga: Hur troligt är det att få 65 eller fler kronor med ett rättvist mynt?

Om  $p$ -värdet  $= 0.003$ , betyder det att denna extrema utfall skulle inträffa i endast 0.3% av fallen med ett rättvist mynt. Vi förkastar därför nollhypotesen och drar slutsatsen att myntet troligen inte är rättvist.

## 4 Felmarginaler (Margin of Error)

### 4.1 Vad är en felmarginal?

Felmarginalen anger det intervall inom vilket det sanna populationsvärdet troligen ligger, baserat på ditt urval.

$$\text{Felmarginal} = Z \times \frac{\sigma}{\sqrt{n}}$$

- $Z$  = Z-värde för önskad konfidensnivå
- $\sigma$  = Standardavvikelse
- $n$  = Urvalsstorlek

### 4.2 Konfidensintervall

Ett konfidensintervall konstrueras som:

$$\text{Estimat} \pm \text{Felmarginal}$$

Exempel: Om 52% av urvalet säger "ja" med felmarginal  $\pm 3\%$ , är konfidensintervallet 49% till 55%.

### 4.3 Faktorer som påverkar felmarginalen

#### 1. Konfidensnivå ( $Z$ )

- Högre konfidensnivå  $\rightarrow$  Större felmarginal
- 90% konfidensintervall är smalare än 95%

#### 2. Urvalsstorlek ( $n$ )

- Större urval  $\rightarrow$  Mindre felmarginal
- Kvadratroten betyder att du behöver 4 gånger så stort urval för att halvera felmarginalen

#### 3. Variation i populationen ( )

- Mer variation  $\rightarrow$  Större felmarginal
- Mer homogen population  $\rightarrow$  Mindre felmarginal

### 4.4 Exempel: Opinionsundersökning

En undersökning av 1000 väljare visar att 48% tänker rösta på parti A.

Med 95% konfidensnivå: -  $Z = 1.96$  - Standardfel  $0.016$  (beräknas från  $p(1-p)/n$ ) - Felmarginal  $= 1.96 \times 0.016 = 0.031 = \pm 3.1\%$

**Tolkning:** Med 95% säkerhet ligger det sanna stödet för parti A mellan 44.9% och 51.1%.

## 5 Samband mellan begreppen

### 5.1 Z-värden och p-värden

Z-värden används för att beräkna p-värden:

| Z-värde | p-värde (tvåsidigt) | Tolkning                                    |
|---------|---------------------|---|
| 1.96    | 0.05                | Gränsen för signifikans vid $\alpha = 0.05$ |
| 2.58    | 0.01                | Starkt signifikant                          |
| 3.29    | 0.001               | Mycket starkt signifikant                   |

### 5.2 Z-värden och felmarginaler

Högre Z-värde (högre konfidensnivå) ger större felmarginal:

| Konfidensnivå | Z     | Effekt på felmarginal |
|---------------|-------|-----------------------|
| 90%           | 1.645 | Smalare intervall     |
| 95%           | 1.96  | Standard              |
| 99%           | 2.576 | Bredare intervall     |

## 6 Praktisk tillämpning

### 6.1 Hypotestest i 5 steg

#### 1. Formulera hypoteser

- $H_0$  : Nollhypotes (ingen effekt/skillnad)
- $H_a$  : Alternativhypotes (det finns en effekt/skillnad)

#### 2. Välj signifikansnivå

- Vanligen  $\alpha = 0.05$

#### 3. Samla data och beräkna teststatistika

- T.ex. Z-värde

#### 4. Beräkna p-värde

- Sannolikheten för observerade data under  $H_0$

#### 5. Dra slutsats

- Om  $p < \alpha$  : Förkasta  $H_0$
- Om  $p \geq \alpha$  : Behåll  $H_0$

## 6.2 Exempel: Är studenter längre än genomsnittet?

### 1. Hypoteser:

- $H_0$  : = 170 cm (studenter är lika långa som genomsnittet)
- $H_1$  : > 170 cm (studenter är längre)

### 2. Signifikansnivå: = 0.05

### 3. Data:

- Urval:  $n = 50$  studenter
- Medellängd:  $\bar{x} = 173$  cm
- Standardavvikelse:  $s = 8$  cm

### 4. Beräkning:

$$z = \frac{173 - 170}{8/\sqrt{50}} = \frac{3}{1.13} = 2.65$$

p-värde 0.004

5. **Slutsats:**  $p < 0.05$ , så vi förkastar  $H_0$ . Det finns statistiskt signifikant bevis för att studenter i genomsnitt är längre än 170 cm.

## 7 Vanliga misstag att undvika

### 7.1 1. P-hacking

Att testa många olika hypoteser tills man hittar en signifikant ( $p < 0.05$ ) leder till falska positiva resultat.

### 7.2 2. Förväxla statistisk och praktisk signifikans

Ett resultat kan vara statistiskt signifikant ( $p < 0.05$ ) men praktiskt oviktigt (mycket liten effekt).

### 7.3 3. Ignorera urvalsstorlek

Stora urval kan ge signifikanta resultat även för små, praktiskt ointressanta effekter.

### 7.4 4. Feltolkning av konfidensintervall

“95% konfidensintervall” betyder INTE att det är 95% chans att det sanna värdet ligger i intervallet. Det betyder att om vi upprepar studien många gånger, kommer 95% av intervallen att innehålla det sanna värdet.

## 8 Viktiga begrepp

| Svenska            | Engelska               | Symbol         |
|--------------------|------------------------|----------------|
| Z-värde            | Z-score                | $z$            |
| P-värde            | P-value                | $p$            |
| Felmarginal        | Margin of error        | $E$ eller $ME$ |
| Konfidensnivå      | Confidence level       | -              |
| Konfidensintervall | Confidence interval    | $CI$           |
| Signifikansnivå    | Significance level     | $\alpha$       |
| Nollhypotes        | Null hypothesis        | $H_0$          |
| Alternativhypotes  | Alternative hypothesis | $H_1$          |
| Standardfel        | Standard error         | $SE$           |

## 9 Interaktiva verktyg

- [Sannolikhet](#) - Utforska normalfördelning och konfidensintervall
- [Urvalsstorlekskalkylator](#) - Se hur Z-värden påverkar nödvändig urvalsstorlek
- [Regression](#) - Inferens för korrelation och regression

## 10 Vidare läsning

För djupare förståelse, se: - [Korrelation och regression](#) - Hypotestest för samband - Presentationer om [Slump och sannolikhet](#)

## 11 Referenser

- Wikipedia: [Standard score \(Z-score\)](#)
- Wikipedia: [P-value](#)
- Wikipedia: [Margin of error](#)
- Wikipedia: [Confidence interval](#)
- Wikipedia: [Statistical hypothesis testing](#)
- Wikipedia: [Significance level](#)
- Wikipedia: [Normal distribution](#)
- Wikipedia: [68–95–99.7 rule](#)