

## HYPOTESPRÖVNING

De statistiska metoderna som används för att fatta denna typ av beslut baseras på två komplementära antaganden om populationen.

Partiets produkter har antingen den utlovade kvaliteten eller så har de den inte.  
Antingen fungerar tillverkningsprocessen som planerat eller så gör den det inte.  
Antingen har behandlingen effekt på patienterna eller så har den det inte.

Dessa antaganden kallas vanligen **hypoteser** och metoderna att fatta beslut av det slag vi har exemplifierat ovan, baserade på observationer från ett urval, kallas **hypotesprövning**.

Vi kommer att diskutera metoder för att pröva hypoteser om populationsparametrar.

### Nollhypotes och mothypotes

Med hjälp av data från ett stickprov vill vi pröva hypoteser om den population eller den sannolikhetsfördelning som stickprovet kommer från.

Vi har en **nollhypotes** ( $H_0$ ) som ställs mot en **mothypotes** ( $H_1$ ).

**Frågan är:** Ger stickprovsdata anledning att förkasta  $H_0$  (till förmån för  $H_1$ ) eller skall vi hålla fast vid  $H_0$ ?

**Ex.:** I en politisk partisympatiundersökning säger sig 14% av de utvalda personerna sympatisera med parti A. Vi vet att i senaste valet fick parti A 12% av rösterna.

**Har andelen A-sympatisörer i populationen ökat sedan valet?**

$H_0$ : Andelen är oförändrad.

$H_1$ : Andelen har ökat.

Skall vi förkasta  $H_0$  eller ej?

Om vi anser att det höga värdet i stickprovet kan förklaras av slumpen (ifall  $H_0$  är sann), så håller vi fast vid  $H_0$ .

Om vi anser att andelen i stickprovet är högre än vad som rimligen kan förklaras av slumpen, så förkastar vi  $H_0$ .

Statistisk hypotesprövning innebär att vi använder vissa **beslutsregler** för när vi skall förkasta  $H_0$  (och när vi skall hålla fast vid  $H_0$ ). Dessa beslutsregler är utformade så att vi har en viss kontroll över risken för felaktigt beslut.

Det finns en asymmetri i behandlingen av  $H_0$  och  $H_1$  i statistisk hypotesprövning.

Som  $H_0$  väljer vi den hypotes som vi i det längsta håller fast vid. Vi kräver extra starkt stöd från observerade data för att  $H_0$  skall förkastas. "Bevisbördan" ligger hos den som förespråkar  $H_1$ .

Ofta innebär  $H_0$  något i stil med "ingen förändring", "ingen effekt", "ingen skillnad", medan  $H_1$  kanske är den ur tillämpningssynpunkt djärvare och mer intressanta hypotesen.

Hypoteser avser oftast värden på populationsparametrar.

Exempel på hypoteser om populationsmedelvärden:

$$H_0: \mu = 50 \quad \text{mot} \quad H_1: \mu \neq 50 \quad (\text{enkel nollhypotes; tvåsidig mothypotes})$$

$$H_0: \mu \leq 50 \quad \text{mot} \quad H_1: \mu > 50 \quad (\text{sammansatt nollhypotes; ensidig mothypotes})$$

Exempel på hypoteser om populationsproportioner:

$$H_0: \pi = 0,1 \quad \text{mot} \quad H_1: \pi \neq 0,1 \quad (\text{enkel nollhypotes; tvåsidig mothypotes})$$

$$H_0: \pi = 0,12 \quad \text{mot} \quad H_1: \pi > 0,12 \quad (\text{enkel nollhypotes; ensidig mothypotes}).$$

### Hypotesprövning (allmänt)

Det gäller att fatta ett beslut: **Förkasta  $H_0$  eller ej?**

Vi kan aldrig vara helt säkra på fatta rätt beslut. Det finns två typer av fel som vi kan göra:

| Verkligheten:<br>Beslut: | $H_0$ sann                               | $H_0$ inte sann                           |
|--------------------------|------------------------------------------|-------------------------------------------|
| Ej förkasta $H_0$        | Riktigt beslut                           | Felaktigt beslut<br>( <b>Typ II-fel</b> ) |
| Förkasta $H_0$           | Felaktigt beslut<br>( <b>Typ I-fel</b> ) | Riktigt beslut                            |

Vid klassisk hypotesprövning håller man sannolikheten för **typ I-fel** under kontroll. Man ser till att den får ett i förväg bestämt värde,  $\alpha$ , som sätts lågt (ofta  $\alpha = 0,01$ ,  $0,05$  eller  $0,10$ ).

$$\alpha = \text{testets signifikansnivå} = P(\text{Förkasta } H_0 \mid H_0 \text{ sann})$$

Tillvägagångssättet vid klassisk hypotesprövning är i huvudsak följande:

1. Formulera *hypoteser*,  $H_0$  och  $H_1$ .
2. Bestäm en *signifikansnivå*  $\alpha$  = sannolikheten (risken) att förkasta  $H_0$ , när  $H_0$  är sann. Ofta väljs  $\alpha = 0,01$ ;  $0,05$  eller  $0,10$ .
3. Ange vilken *testvariabel* som skall användas. **Testvariabel** = en storhet vars värde skall beräknas från stickprovsdata. Detta värde skall sedan ligga till grund för vårt beslut.
4. Ange *beslutsregel*. Dvs ange förkastelsegränser, sådana att  $H_0$  skall förkastas om testvariabeln antar ett värde utanför dessa gränser. (Gränserna skall alltså bestämmas så att testvariabeln med just sannolikheten  $\alpha$  kommer att anta ett värde utanför dessa, ifall  $H_0$  är sann.)
5. **Beräkna** testvariabelns värde med användning av erhållna stickprovsdata.
6. **Slutsats**. Om testvariabeln antar ett värde *utanför* förkastelsegränserna, så **förkastas  $H_0$**  och vi säger att vi har fått ett resultat som är **signifikant** på signifikansnivån  $\alpha$ .

I annat fall säger vi att  $H_0$  **inte kan förkastas** (ett icke-signifikant resultat).

**OBS!** Ett icke-signifikant resultat innebär inte att vi kan dra slutsatsen att  $H_0$  är sann. Det innebär bara att  $H_1$  inte är någon stark konkurrent till  $H_0$ . Det kan finnas många andra tänkbara nollhypoteser som inte heller skulle ha förkastats.

Att ”icke-förkasta”  $H_0$  är alltså inte detsamma som att utan vidare acceptera  $H_0$ . Kursbokens författare säger:

## 1. Hypotesprövning för ett medelvärde

Antag först att vi har ett stickprov av storlek  $n$  från en **normalfördelad population** med **okänt medel-värde**  $\mu$ , men med **känd varians**  $\sigma^2$ . Vi vill pröva

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Signifikansnivå:  $\alpha = 0,05$ .

$$\text{Testvariabel: } Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

**Beslutsregel:**  $H_0$  förkastas, om  $|Z_{\text{obs}}| > 1,96$  (dvs. om  $Z_{\text{obs}} > 1,96$ , eller  $Z_{\text{obs}} < -1,96$ ).

Med signifikansnivån  $\alpha = 0,01$  förkastas  $H_0$  ifall  $|Z_{\text{obs}}| > 2,576$ .

Med signifikansnivån  $\alpha = 0,10$  förkastas  $H_0$  ifall  $|Z_{\text{obs}}| > 1,645$ .

**Motivering:** En rimlig utgångspunkt är att  $H_0$  skall förkastas ifall vi får ett värde på  $\bar{x}$  som ligger ”långt” ifrån det under  $H_0$  förväntade värdet  $\mu_0$ .

Att  $\bar{x}$  ligger långt ifrån  $\mu_0$  är detsamma som att testvariabeln  $Z$  antar ett värde långt ifrån värdet 0 (antingen åt det positiva eller det negativa hållet).  $H_0$  skall alltså förkastas ifall vi får ett osannolikt högt eller osannolikt lågt värde på  $Z$ .

Vi skall med andra ord förkasta  $H_0$ , ifall vi får ett värde på  $Z$ , som ligger utanför gränserna  $\pm c$ , där  $c$  är en positiv konstant.

Bestäm nu  $c$  så att signifikansnivån blir den önskade ( $\alpha = 0,05$ ), dvs. bestäm  $c$  så att

$$P(|Z| > c \mid H_0 \text{ sann}) = 0,05$$

Vi vet att om  $H_0$  är sann, så har testvariabeln  $Z$  en standardiserad normalfördelning,  $N(0; 1)$ . Då är

$$P(|Z| > 1,96 \mid H_0 \text{ sann}) = 0,05$$

Sätt alltså  $c = 1,96$ . Beslutsregeln blir: förkasta  $H_0$ , om vi får ett  $Z$ -värde utanför gränserna  $\pm 1,96$ .

## Grundprincipen vid statistisk hypotesprövning är:

Om vi får ett värde på testvariabeln, som skulle vara mycket osannolikt ifall  $H_0$  vore sann, så väljer vi att i stället tro på  $H_1$ .

Hur gör man vid **ensidig** mothypotes? T.ex.

$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

Samma testvariabel, men en annan förkastelsegräns.

$H_0$  förkastas (på signifikansnivån 0,05) ifall  $Z_{\text{obs}} > 1,645$ . Det är ju *höga* värden på  $\bar{x}$ , dvs. *höga* värden på  $Z$ , som ger anledning att förkasta  $H_0$  till förmån för  $H_1$ .

Om mothypotesen i stället är  $H_1: \mu < \mu_0$ ,

så förkastas  $H_0$  (på signifikansnivån 0,05) ifall  $Z_{\text{obs}} < -1,645$ .

Hur gör man när populationsvariansen  $\sigma^2$  inte är känd, eller när stickprovet är litet?

**Testvariabler vid test av hypotes om  $\mu$ :**

|                                                                     |                                                                                                                                                  |
|---------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| $n \geq 30$<br>(oavsett om populationen är normalfördelad eller ej) | $\sigma^2$ känd: $Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$<br><br>$\sigma^2$ okänd: $Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ |
| $n < 30$<br>Populationen normalfördelad.                            | $\sigma^2$ känd: $Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$<br><br>$\sigma^2$ okänd: $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ |

När  $H_0$  är sann, så har testvariabeln  $Z$  en standar-diserad normalfördelning (exakt eller approxima-tivt), och testvariabeln  $t$  har en  $t$ -fördelning med  $n-1$  frihetsgrader.

**Exempel 1:** 16 mjölkpaket väljs slumpmässigt bland dem som fyllts på efter justeringen och man beräknar stickprovsmedelvärdet till 1,008 liter. Standardavvikelsen antas inte ha påverkats av justeringen utan man utgår från att  $\sigma=0.02$ .

Hypoteser:  $H_0: \mu = 1$

$H_1: \mu \neq 1$

Sign.-nivå:  $\alpha = 0,05$

Testvariabel:  $Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

Beslutsregel:  $H_0$  förkastas om  $|Z_{\text{obs}}| > 1,96$ .

Resultat:  $Z_{\text{obs}} = \frac{1,008 - 1}{\frac{0,02}{\sqrt{16}}} = 1,60$

Slutsats:  $H_0$  kan inte förkastas på 5% signifikansnivå.

**Exempel 2:** Stickprov på  $n = 70$  observationer från population med okänd fördelning och med okänd standardavvikelse.

Stickprovets medelvärde:  $\bar{x} = 14\,100$

Stickprovets standardavvikelse:  $s = 1\,900$

Hypoteser:  $H_0: \mu = 13\,500$

$H_1: \mu > 13\,500$  (Ensidig mothypotes)

Sign.-nivå:  $\alpha = 0,05$

Testvariabel:  $Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Beslutsregel:  $H_0$  förkastas om  $Z_{\text{obs}} > 1,64$ .

Resultat:  $Z_{\text{obs}} = \frac{14\,100 - 13\,500}{\frac{1\,900}{\sqrt{70}}} = 2,64 > 1,64$

Slutsats:  $H_0$  förkastas på 5% signifikansnivå.

**Exempel 3:** Antag att vi vill pröva hypotesen att Martin Martini's restid till arbetet är 30 minuter. Standardavvikelsen i Martin Martini's restid är okänd. Ett urval om 20 observationer dras. I urvalet uppskattas standardavvikelsen till 2, stickprovets medelvärde  $\bar{x} = 26,5$ .

Hypoteser:  $H_0: \mu = 30$  (Sätt  $\mu_0 = 30$ )

$H_1: \mu \neq 30$

Sign.-nivå:  $\alpha = 0,05$

Testvariabel:  $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Frihetsgrader:  $n - 1 = 19$

Beslutsregel:  $H_0$  förkastas om  $|t_{\text{obs}}| > 2,093$

Resultat:  $t_{\text{obs}} = \frac{26,5 - 30}{\frac{2}{\sqrt{20}}} = 0,94$

Slutsats:  $H_0$  kan inte förkastas på 5% signifikansnivå.

**Exempel 4:** En cigarettillverkare hävdar i sin reklam att de cigaretter de tillverkar innehåller mindre än 25 mg nikotin i genomsnitt. Ett urval om 36 cigaretter gav medelvärdet 24,5 mg och standardavvikelsen 3 mg.

Vi prövar nu cigarettillverkarens påstående genom att pröva hypoteser:

$$H_0: \mu = 25$$

$$H_1: \mu < 25 \quad (\text{Ensidig mothyp.})$$

Sign.-nivå:  $\alpha = 0,05$

Testvariabel:  $Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Beslutsregel:  $H_0$  förkastas om  $Z_{\text{obs}} < -1,64$ .

Resultat:  $Z_{\text{obs}} = \frac{24,5 - 25}{\frac{3}{\sqrt{36}}} = -1 > -1,64$

Slutsats:  $H_0$  förkastas ej på 5% signifikansnivå.

## 2. Hypotesprövning för en differens mellan två medelvärden

|              | Medelvärde | Varians    |
|--------------|------------|------------|
| Population 1 | $\mu_X$    | $\sigma_X$ |
| Population 2 | $\mu_Y$    | $\sigma_Y$ |

Vi är intresserade av differensen  $\mu_X - \mu_Y$ .

|                          | Antal obs.<br>i stickpr. | Medel-<br>värde | Varians |
|--------------------------|--------------------------|-----------------|---------|
| Stickprov<br>från pop. 1 | $n_X$                    | $\bar{x}$       | $s_X$   |
| Stickprov<br>från pop. 2 | $n_Y$                    | $\bar{y}$       | $s_Y$   |

**Oberoende antas mellan de båda stickproven.**

Hypoteser:

$$H_0: \mu_X - \mu_Y = D_0$$

$$H_1: \mu_X - \mu_Y \neq D_0 \text{ (eller enkelsidig mothyp.)}$$

där  $D_0$  är ett tal.

Ofta är  $D_0 = 0$ , dvs. vi prövar nollhypotesen:

$$H_0: \mu_X - \mu_Y = 0$$

Vi vill med andra ord pröva om båda populationer-na kan tänkas ha samma medelvärde.

Vi vet att  $\bar{x} - \bar{y}$  är en väntevärdesriktig skattning av  $\mu_X - \mu_Y$ .

Frågan vid hypotesprövningen är den vanliga: Ligger det observerade värdet på  $\bar{x} - \bar{y}$  så långt ifrån talet  $D_0$ , att vi finner anledning att förkasta  $H_0$ ?

Här krävs att båda stickproven kommer från populationer som är **normalfördelade**.

Testvariabel när populationsvarianserna är **kända**:

$$Z = \frac{(\bar{x} - \bar{y}) - D_0}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \stackrel{H_0}{\sim} N(0; 1)$$

När populationsvarianserna är **okända** krävs dessutom att dessa är **lika stora**. (känner inte till det numeriska värdet på  $\sigma^2$ .)

Testvariabel:

$$t = \frac{(\bar{x} - \bar{y}) - D_0}{\sqrt{s_P^2 \left( \frac{1}{n_X} + \frac{1}{n_Y} \right)}} \stackrel{H_0}{\sim} t(n_X + n_Y - 2)$$

$$S_P^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}$$

Om observationerna är oberoende,  $n_X > 30$ ,  $n_Y > 30$  (stora stickprov) men med okänd fördelning får vi lita till centrala gränsvärdessatsen.

Testvariabel när populationsvarianserna är **kända**:

$$Z = \frac{(\bar{x} - \bar{y}) - D_0}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

Testvariabel när populationsvarianserna är **okända**:

$$Z = \frac{(\bar{x} - \bar{y}) - D_0}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}}$$

Testvariablerna  $Z$  ovan är approximativt normal-fördelade,  $N(0; 1)$ , när  $H_0$  är sann.

**Exempel.** Vid en industri finns en maskin A som producerar vissa enheter. Man funderar över att köpa en ny maskin, B, som visserligen har en högre produktionshastighet när den fungerar som den ska, men också ett större antal driftstopp. Man beslutar att köpa maskin B om dess genomsnittsproduktion är högre än maskin A:s genomsnittsproduktion. Båda maskinerna provas under 100 dygn och man erhåller

|                     | <b>Maskin A</b> | <b>Maskin B</b> |
|---------------------|-----------------|-----------------|
| Antal observationer | 100             | 100             |
| Medelproduktion     | 1305            | 1338            |
| Standardavvikelse   | 25              | 100             |

Hypoteserna formuleras som

$$H_0 : \mu_A - \mu_B = 0$$

$$H_1 : \mu_A - \mu_B < 0$$

Om maskin B genomsnittligt sett är bättre än A blir differensen  $\mu_A - \mu_B$  negativ, vilket anges i mothypotesen, som på grund av frågeställningen är ensidig.

$$Z_{OBS} = \frac{(1305 - 1338) - 0}{\sqrt{\frac{25^2}{100} + \frac{100^2}{100}}} = \frac{-33}{\sqrt{106,25}} = \frac{-33}{10,308} = -3,2$$

$Z_{OBS} < -1,64 \Rightarrow H_0$  förkastas på 5% sign.-nivå, dvs maskin B är bättre.



### 3. Prövning av hypoteser om proportioner

Att pröva en hypotes om proportioner bygger på att binomialfördelningen kan approximeras med normalfördelningen.

Antag t ex att andelen individer med en viss egenskap i en stor population är  $\pi$ .

Låt  $X$  vara antalet individer med den egenskapen i ett urval om  $n$  individer.  $X \sim \text{Bin}(n; \pi)$ .

$E[X] = n\pi$  och  $V[X] = n\pi(1 - \pi)$ .

Låt proportionen av individer i urvalet med den studerade egenskapen vara  $\hat{P}$ .

$$\hat{P} = X / n$$

$E[\hat{P}] = \pi$  och  $V[\hat{P}] = \frac{\pi(1-\pi)}{n}$ .

Om urvalet är stort

$$Z = \frac{\hat{P} - \pi}{\sqrt{\pi(1-\pi)/n}} \underset{\text{approx}}{\sim} N(0;1)$$

**Exempel.** För en stor population av kvinnor i en viss ålder gäller vid en viss tidpunkt att 64 % är körkortsinnehavare, vilket man fastställt genom en totalundersökning. Vid ett senare tillfälle vill man undersöka om andelen körkortsinnehavare i den aktuella åldern är oförändrad. Man drar därför ett urval om 200 individer, varav 160 visar sig vara körkortsinnehavare. Vi vill nu pröva på signifikansnivån 5 % om andelen körkortsinnehavare i populationen har ändrats.

Hypoteserna blir

$$H_0 : \pi = 0,64$$

$$H_A : \pi \neq 0,64$$

Nollhypotesen anger här att andelen körkortsinnehavare är oförändrad medan mothypotesen säger att en förändring har ägt rum. Alternativhypotesen är således tvåsidig.

En förändring kan ju antingen bestå i att andelen har minskat eller att den har ökat.

Om  $H_0$  är sann

$$Z = \frac{\hat{P} - \pi}{\sqrt{\pi(1-\pi)/n}} \underset{H_0}{\sim} N(0;1)$$

$$Z_{OBS} = \frac{0,8 - 0,64}{\sqrt{0,64(1 - 0,64)/200}} = \frac{0,16}{0,034} = 4,71$$

Beslutregel:  $|Z_{OBS}| > 1,96$  förkasta  $H_0$  på signifikansnivån 5%.

$Z_{OBS} = 4,71 > 1,96$ , dvs hypotesen om ingen förändring förkastas på signifikansnivån 5%.