© 2005 Eric Järpe
Högskolan i Halmstad

Hypotestest

Statistiska test används för att urskilja systematiska fel från slumpmässig avvikelse.

Exempel På en arbetsplats sätts lönerna efter hur stor arbetsprestationen är. Arbetarna presterar per dag under en arbetsvecka

$$60 \quad 66 \quad 63 \quad 55 \quad 63 \quad 58$$

enheter. De tycker att lönen är för låg men arbetsgivaren hävdar att i det långa loppet ligger genomsnittet på 57.9 enheter och att lönen därför är motiverad. Arbetarna tror att att deras presation är större. Under antagandet att presationen är fördelad med standardavvikelse 4.35, har arbetarna rätt till högre lön?

Lösning: För att resonera sig fram till en lösning av detta problem, utgår vi från det vi känner till om den vvr skattningen av μ: $\bar{X}$ och dess approximativa fördelning $N(\mu, \frac{\sigma}{\sqrt{n}})$.

Låt X vara antalet enheter som produceras en slumpmässigt vald dag. Under antagandet att arbetsgivaren har rätt är $\mu = 57.9$ varmed $\bar{X}$ skulle vara fördelat $N(57.9, \frac{4.35}{\sqrt{6}}) = N(57.9, 1.77588)$. Nu har vi emellertid ett stickprov från vilket vi kan beräkna det *observerade* medelvärdet $\bar{x} = 60.8333$. Om nu detta är så stort att sannolikheten att vi skulle få ett minst så stort värde är mindre än en viss föreskriven felmarginal kan vi anse oss har argument för att arbetarna har rätt i att den förväntade prestationen är större än 57.9.

Detta är i korta ordalag logiken bakom ett statistiskt test. Man utgår från den fördelning man *inte* tror på och ser om fördelningsantagandet är orimligt med tanke på vad man observerat (se figur 9.2 i Vännman).

I detta exempel får vi följande resultat. Antag att vi väljer felmarginalen 5%. Detta är då risken att μ trots allt faktiskt är 57.9 trots att vi fick ett så stort medelvärde från stickprovet. Dvs vi vill välja ett x-värde c som är så stort att $P(\bar{X} > c|\mu = 57.9) = 0.05$, och sedan jämföra vår medelvärdesobservation $\bar{x}$ med c. Vi får: $0.05 = P(\bar{X} > c|\mu = 57.9) = 1 - P(\bar{X} \leq c|\mu = 57.9) = 1 - \Phi(\frac{c-57.9}{4.35/\sqrt{6}}) \Rightarrow$
$\Rightarrow \Phi(\frac{c-57.9}{4.35/\sqrt{6}}) = 0.95 \Rightarrow \frac{c-57.9}{4.35/\sqrt{6}} = \lambda_{0.05} = 1.644854 \Rightarrow c = 1.644854 \cdot \frac{4.35}{\sqrt{6}} + 57.9 =$
$= 60.821$. Detta är alltså gränsen för vad vi som är "5%-rimligt" − om vi får ett medelvärde under den finns det ingen anledning att betvivla att väntevärdet är 57.9 eller lägre, men om vi får ett värde över denna gräns tror vi att $\mu > 57.9$. I vårt fall hade vi att $\bar{x} = 60.8333 > 60.821$ så vi klarar oss med nöd och näppe, dvs *vi har bevisat på 5% signifikansnivå att arbetsgivaren har fel.*

Om medelvärdet hade varit mindre än c hade vi inte kunnat bevisa någonting! Dvs detta hade *inte* bevisat att $\mu = 57.9$. □

Visserligen beskriver exemplet ovan precis förfarandet vid hypotestest men vi vill kunna göra många olika sorters test och ett sådant resonemang för varje specialfall blir för omständigt. Därför finns en mer systematisk procedur för hur man går till väga då man genomför ett test, men den bakomliggande logiken är alltid densamma.

Arbetsgång vid hypotestest

1. Ställ upp en hypotes: $\begin{cases} H_0 : \theta = \theta_0 & \textbf{nollhypotes} \\ H_1 : \theta \in \Theta & \textbf{mothypotes} \end{cases}$

2. Observera ett stickprov $x_1, \ldots, x_n$.
   (Motsvarande stokastiska variabler $X_1, \ldots, X_n$ har en fördelning F_X under H_0.)

3. Välj signifikansnivå α (felmarginalen).

4. Beräkna teststatistikan $u = u(x_1, \ldots, x_n)$.
   (Motsvarande stokastisk variabel $U = U(X_1, \ldots, X_n)$ har fördelning F_U under H_0).

5. Beräkna kritiskt område C_α för U enligt F_U och α.
   (Dvs C_α ska uppfylla villkoret $P(U \in C_\alpha | H_0) = \alpha$).

6. Testregel:
   - Om $u \in C_\alpha$, förkasta H_0
     (då har man bevisat H_1 med $100(1-\alpha)\%$ säkerhet).
   - Om $u \notin C_\alpha$, förkasta *ej* H_0
     (då kan man *ej* visa något!).

Tilläggas kan att *den minsta signifikansnivå som en hypotes kan förkastas med* kallas ***p*-värdet**, t.ex. om mothypotesen är $H_1 : \mu > \mu_0$ och således $C_\alpha = \{u : u > \lambda_\alpha\}$ så är p-värdet $P(U > u | H_0)$ (och motsvarande med andra hypoteser).

Test av μ då σ är känt

1. Stickprovet
2. Hypotes A: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ B: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ C: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$
3. Signifikansnivå α
4. Teststatistika $U = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \in N(0,1)$ under H_0
5. Kritiskt område
   A: $C_\alpha = \{u : u < -\lambda_\alpha\}$ B: $C_\alpha = \{u : u > \lambda_\alpha\}$ C: $C_\alpha = \{u : |u| > \lambda_{\alpha/2}\}$
6. Förkasta om $u \in C_\alpha$, inte annars!

Test av μ då σ är okänt

1. Stickprovet
2. Hypotes A: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ B: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ C: $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$
3. Signifikansnivå α
4. Teststatistika $U = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \in t(n-1)$ under H_0
5. Kritiskt område A: $C_\alpha = \{u : u < -t_\alpha(n-1)\}$
   B: $C_\alpha = \{u : u > t_\alpha(n-1)\}$ C: $C_\alpha = \{u : |u| > t_{\alpha/2}(n-1)\}$
6. Förkasta om $u \in C_\alpha$, inte annars!

Test av σ^2

1. Stickprovet
2. Hypotes A: $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases}$ B: $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases}$ C: $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$
3. Signifikansnivå α
4. Teststatistika $U = \dfrac{(n-1)S^2}{\sigma_0^2} \in \chi^2(n-1)$ under H_0
5. Kritiskt område A: $C_\alpha = \{u : u < \chi^2_{1-\alpha}(n-1)\}$ B: $C_\alpha = \{u : u > \chi^2_\alpha(n-1)\}$
   C: $C_\alpha = \{u : u < \chi^2_{1-\alpha/2}(n-1)$ eller $u > \chi^2_{\alpha/2}(n-1)\}$
6. Förkasta om $u \in C_\alpha$, inte annars!

Låt oss återbesöka förlossningsexemplet.

Exempel Tiden för förlossning varierar mellan olika sjukhus och olika delar av världen. Vid ett sjukhus har den genomsnittliga förlossningstiden $\bar{x}$ beräknats baserat på 5 obervationer. Vid en landsomfattande studie har man funnit att den förväntade förlossningstiden är 14 timmar och standardavvikelsen är 1. Kan man med 5% felmarginal säga att en (ännu ej observerad) tid från sjukhuset har

(a) högre väntevärde än i övriga landet om $\bar{x} = 14.84$ och $\sigma = 1$?

(b) lägre väntevärde än i övriga landet om $\bar{x} = 12.84$ och σ är okänd men $s = 1.3$?

(c) större varians än i övriga landet ($\sigma^2 = 1$) om $s = 1.3$?

Lösning:

(a) Hypotes: $\begin{cases} H_0 : \mu = 14 \\ H_1 : \mu > 14 \end{cases}$

Signifikansnivå $\alpha = 0.05$.

Teststatistika $u = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{14.84-14}{1/\sqrt{5}} = 1.88$

Kritiskt område $C_{0.05} = \{u : u > \lambda_{0.05}\} = (1.64, \infty)$.

Utslag: H_0 förkastas eftersom $u = 1.88 \in (1.64, \infty)$, dvs man har på 5% signifikansnivå visat att sjukhuset har högre förväntad förlossningstid än i övriga landet.

(b) Hypotes: $\begin{cases} H_0 : \mu = 14 \\ H_1 : \mu < 14 \end{cases}$

Signifikansnivå $\alpha = 0.05$.

Teststatistika $u = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{12.84-14}{1.3/\sqrt{5}} = -1.995$

Kritiskt område $C_{0.05} = \{u : u < -t_{0.05}(5-1)\} = (-\infty, -2.1318)$.

Utslag: H_0 kan ej förkastas eftersom $u = -1.995 \notin (-\infty, -2.1318)$, dvs man kan på signifikansnivå 5% ej visa något.

(c) Hypotes: $\begin{cases} H_0 : \sigma^2 = 1 \\ H_1 : \sigma^2 > 1 \end{cases}$

Signifikansnivå $\alpha = 0.05$.

Teststatistika $u = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(5-1)1.3^2}{1} = 6.76$

Kritiskt område $C_{0.05} = \{u : u > \chi^2_{0.05}(5-1)\} = (9.4877, \infty)$.

Utslag: H_0 kan ej förkastas eftersom $u = 6.76 \notin (9.4877, \infty)$, dvs man kan på signifikansnivå 5% ej visa något. $\square$

Om man inte misstänker att $\mu > \mu_0$ eller $\mu < \mu_0$, *innan* man observerat stickprovet, är den naturliga mothypotesen $H_1 : \mu \neq \mu_0$.

Exempel Antag att vi via undersökningar vet att den förväntade längden av flickor i åldrarna 20 till 30 år är 165 cm. Man vill veta om fotomodeller skiljer sig (i fråga om längd) från den övriga befolkningen och mäter därför 10 slumpvis valda modeller med resultatet $\sum x_i = 1403$ och $\sum x_i^2 = 203\,000$. Kan man bevisa något på signifikansnivå 0.05% från dessa observationer?

Lösning: Låt X vara längden av en modell och $\mu = E(X)$. Vi vill testa hypotesen

$$\begin{cases} H_0 : \mu = 165 \\ H_1 : \mu \neq 165 \end{cases}$$

Från $\sum x_i = 1403$ får vi att $\bar{x} = 140.3$ och från $\sum x_i^2 = 203\,000$ fås att $s^2 = 684.34$ varmed $u = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{-14.7}{\sqrt{68.434}} = -2.986$. Vidare har vi att $t_{0.05/2}(10-1) = 2.2622$ och eftersom $|u| = 2.986 > 2.2622 = t_{\alpha/2}(n-1)$ så förkastas H_0, dvs ja, modellerna skiljer sig på signifikansnivå 5% från befolkningen i övrigt beträffande längd. $\square$