Sannolikhetslära och Statistik
Föreläsning 9

Mattias Villani

Avdelningen för Statistik och Maskininlärning
Institutionen för datavetenskap
Linköpings universitet

ÖVERSIKT

- **Hypotesttest**
- **Chi-tvåfördelningen** (χ^2)
- χ^2**-test modellutvärdering**

HYPOTESTEST

- Ex 1. genomsnittshastigheten på ditt bredband är sämre än leverantören utlovat.

$$\textbf{Nollhypotes}: \; H_0 : \mu \geq 8\text{Mbit/s}$$

$$\textbf{Alternativhypotes}: \; H_A : \mu < 8\text{Mbit/s}$$

- Ex 2. En ny medicin påverkar blodtrycket.

$$\textbf{Nollhypotes}: \; H_0 : \mu = 0$$

$$\textbf{Alternativhypotes}: \; H_A : \mu \neq 0$$

- Ex 3. en UI-förändring ökar andelen nöjda användare.

$$\textbf{Nollhypotes}: \; H_0 : p \leq p_0$$

$$\textbf{Alternativhypotes}: \; H_A : p > p_0$$

- Ex 4. andelen KD-väljare är under 4%-spärren.

HYPOTESTEST

- **Tvåsidigt test** förkastar H_0 om μ är större **eller** mindre än μ_0

$$\textbf{Nollhypotes}: \; H_0 : \mu = \mu_0$$

$$\textbf{Alternativhypotes}: \; H_A : \mu \neq \mu_0$$

- **Ensidigt test**

$$\textbf{Nollhypotes}: \; H_0 : \mu \leq \mu_0$$

$$\textbf{Alternativhypotes}: \; H_A : \mu > \mu_0$$

eller

$$\textbf{Nollhypotes}: \; H_0 : \mu \geq \mu_0$$

$$\textbf{Alternativhypotes}: \; H_A : \mu < \mu_0$$

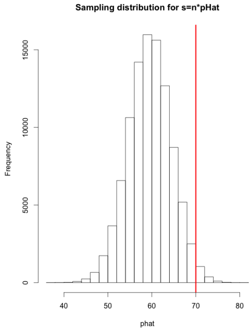
- Ensidiga test skrivs ibland så här (det ger samma resultat):

$$\textbf{Nollhypotes}: \; H_0 : \mu = \mu_0$$

$$\textbf{Alternativhypotes}: \; H_A : \mu < \mu_0$$

HYPOTESTEST

- Notera att **hypotesen gäller populationen** (μ eller p). För en patient kan förändring i blodtrycket vara skild från noll, trots att $\mu = 0$.
- **Stickprov** från populationen. Dra **slutsatser** om H_0 eller H_A är sann.
- Ex 3. $p_0 = 0.6$. $s = 70$ baserat på ett stickprov av $n = 100$ användare. Är $s = 70$ **tillräckligt stort** för att förkasta $H_0 : p \leq p_0$?
- Prova: `sum(runif(100)<=0.6)` i R.



Hypotestest

- **Typ I fel**

$$\alpha = P\left\{\text{Förkasta } H_0 \middle| H_0 \text{ är sann}\right\}$$

  Vi vill kontrollera att α hålls på en förbestämd låg nivå.

  - Ex Gödsel inte effektivt, men du väljer att gödsla ändå.

- **Typ II fel**

$$P\left\{\text{Acceptera } H_0 \middle| H_0 \text{ är falsk}\right\}$$

  - Ex Gödsel effektivt, men du väljer att inte gödsla.

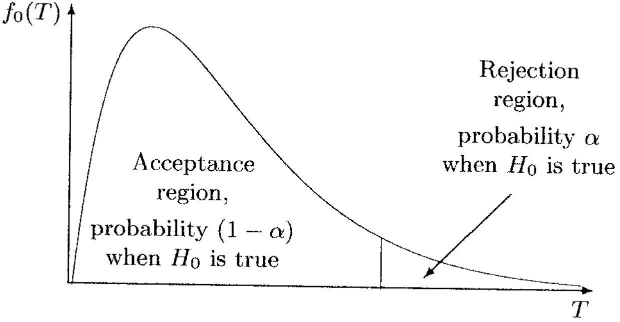
	Acceptera H_0	Förkasta H_0
H_0 sann	Korrekt beslut	Typ 1 fel
H_A sann	Typ II fel	Korrekt beslut

- Popper: man kan bara förkasta en nollhypotes, aldrig acceptera den.
- **Styrka** (power): $p(\theta) = P\left\{\text{Förkasta } H_0 \middle| H_A \text{ är sann}\right\}$.
- Applet: http://tube.geogebra.org/student/m137287.

Steg vid hypotestest

1. Välj **teststatistiska**, $T = T(X_1, ..., X_n)$.
2. Beräkna **nollfördelningen** för T (dvs samplingfördelningen F_0 för T om H_0 är sann).
3. Bestäm **förkastningsregionen** $\mathcal{R}$ i nollfördelningen så att $P\{T \in \mathcal{R}\} = \alpha$.
4. Förkasta H_0 på **signifikansnivån** α om $T_{obs} \in \mathcal{R}$, där T_{obs} är det **observerade** värdet på T.

FÖRKASTNINGSREGIONEN



Steg vid hypotestest - Bernoulliexempel

1. **Teststatistiska**, $T = S = \sum_{i=1}^{n} X_i = n\hat{p}$.

2. **Nollfördelningen** för T: $\sum_{i=1}^{n} X_i \sim Bin(n, p_0)$.

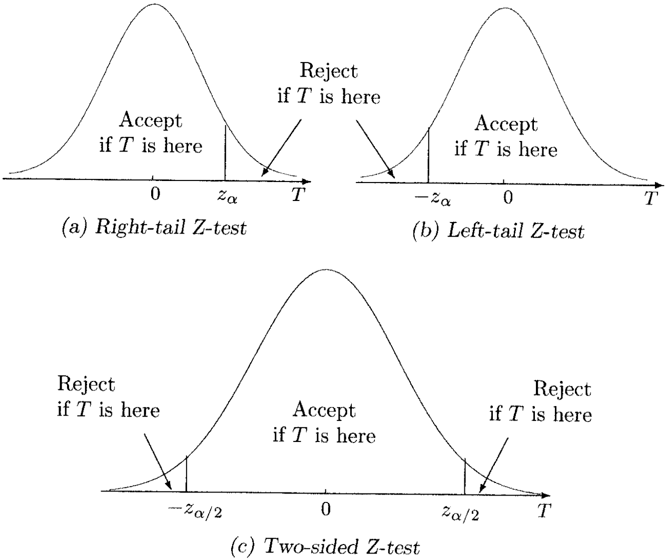
3. Låt $\alpha = 0.05$. `qbinom(p=0.05,size=100,prob=0.6, lower.tail=F)` ger $\mathcal{R} = [68, 100]$ (ungefär).

4. $T_{obs} = 70$ så $T_{obs} \in \mathcal{R}$ och nollhypotesen $p \leq 0.6$ förkastas på signifikansnivån 0.05.

Z-TEST

- **Z-test** används när nollfördelningen är normalfördelad.
  - Ex 1. $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ och
$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$
  - Ex 2. CLT.
- Exempel: $H_0 : \mu = \mu_0$ och $H_A : \mu > \mu_0$.
- En-sidigt Z-test
$$\begin{cases} \text{Förkasta } H_0 & \text{om } Z \geq z_\alpha \\ \text{Acceptera } H_1 & \text{om } Z < z_\alpha \end{cases}$$
- Två-sidigt Z-test. $H_0 : \mu = \mu_0$ och $H_A : \mu \neq \mu_0$.
$$\begin{cases} \text{Förkasta } H_0 & \text{om } |Z| \geq z_{\alpha/2} \\ \text{Acceptera } H_1 & \text{om } |Z| < z_{\alpha/2} \end{cases}$$

FÖRKASTNINGSREGIONER - Z-TEST



T-TEST

- **Z-test** används när nollfördelningen är normalfördelad.
- Om $\hat{\theta}$ normalfördelad, men σ^2 **inte** är känd utan skattas med s^2 blir inte
$$\frac{\hat{\theta} - \theta_0}{s/\sqrt{n}}$$
längre normalfördelad utan t-**fördelad** med $n-1$ frihetsgrader.
- z_α blir istället t_α och hämtas från Tabell 5 i Byron. Se avsnitt 9.4.8 i Baron.

Z-test för skillanden mellan populationer

- Vi kan också testa om **två** populationer har samma väntevärde:

$$H_0 : \mu_X = \mu_Y$$

$$H_A : \mu_X \neq \mu_Y$$

- Ex är andelen KD-sympatisörer lika stor i Stockholm och Göteborg?
- Låt $X_1, ...X_n$ vara ett slumpmässigt stickprov från $N(\mu_X, \sigma^2)$.
- Låt $Y_1, ..., Y_m$ vara ett slumpmässigt stickprov från $N(\mu_Y, \sigma^2)$.
- **Teststatistika**: $\bar{X} - \bar{Y}$. Samplingfördelning under H_0?
  - Linjärkombination av normalvariabler är normalfördelade. $\bar{X} - \bar{Y}$ är normalfördelad.
  - $E(\bar{X} - \bar{Y}) = \mu_X - \mu_Y = 0$ under H_0.
  - $Var(\bar{X} - \bar{Y}) = Var(\bar{X}) + Var(\bar{Y}) = \sigma^2/n + \sigma^2/m = \sigma^2 (1/n + 1/m)$.

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma^2 \left(\frac{1}{n} + \frac{1}{m}\right)}}.$$

Koppling mellan hypotestest och konfidensintervall

Ett test av $H_0 : \theta = \theta_0$ mot $H_A : \theta \neq \theta_0$ på signfikansnivån α

accepterar nollhypotesen

om och endast om

θ_0 ingår i ett symmetriskt $(1 - \alpha)100\%$ konfidensintervall för θ

P-VÄRDE

- Hur väljer vi α?
- **Lågt** α ställer mycket **stora krav** på bevisningen: teststatistikan måste anta mycket stora (positiva eller negativa) värden för att vi ska kunna förkasta H_0.
- **Stort** α ställer **låga krav**. Vi förkastar baserat på väldigt lite bevis.
- Idé: presentera resultat för alla α.
- **P-värde** = den lägsta signifikansnivån α där vi kan förkasta nollhypotesen.
- Alternativ definition: Sannolikheten att acceptera en teststatistiska som är lika extrem eller ännu mer extrem än T_{obs}.
- Exempel: ensidigt Z-test:

$$p\text{-värde:} \quad P\left\{Z \geq Z_{obs}\right\}$$

Skatta en varians - χ^2-fördelningen

- Väntevärdesriktig estimator av σ^2

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

- $\chi 2$ **(Chi-två) fördelningen** med ν frihetsgrader

$$f(X) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, \quad x > 0$$

- Om $X \sim \chi^2_\nu$ så gäller

$$\mathbb{E}X = \nu, \qquad Var(X) = 2\nu$$

- $\chi^2_\nu = Gamma(\nu/2, 1/2)$.
- **Samplingfördelning** för σ^2 (om $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$)

$$\frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}$$

- χ^2-fördelningen kan användas för att skapa konfidensintervall och hypotestest för σ^2.

χ^2 GOODNESS OF FIT TEST

- Antag att din population har följande diskreta fördelning F_0 : $P(X = 1) = p_1, P(X = 2) = p_2, ..., P(X = m) = p_m$.
- Om du har observerat n observationer så förväntar du dig np_k observationer där $X = k$.
- Låt $Exp(k)$ beteckna förväntat antal observationer med värde k om F_0 är en korrekt populationsmodell.
- Låt $Obs(k)$ beteckna faktiskt antal observationer med värde k.

χ^2 GOODNESS OF FIT TEST

- **Chi-två statistiskan**

$$\chi^2 = \sum_{k=1}^{N} \frac{[Obs(k) - Exp(k)]^2}{Exp(k)}$$

- Om χ^2 är för stort så drar vi slutsatsen att data inte kommer från populationen med fördelningen F_0 ovan.
- Men hur stort är för stort? Jämför med samplingfördelningen för χ^2 under $H_0 : F = F_0$ mot $H_A : F \neq F_0$.
- Vid stora stickprov följer Chi-två statistikan en χ^2-fördelning med $m - 1$ frihetsgrader, om $Exp(k) > 5$ för alla k.
- Kan även testa om data kommer från $F_0(\theta)$ där θ är en okänd parameter som skattas med en konsistent estimator. Frihetsgrader $= n - 1 - d$.
- **Kontinuerliga fördelningar** kan hanteras genom diskretisering (men se till att $Exp(k) > 5$ för alla k.