TDAB01 Sannolikhetslära och Statistik

Jose M. Peña
IDA, Linköpings Universitet

Föreläsning 9

Översikt

- **Hypotesttest**
- **Z-test**
- **T-test**
- χ^2**-test för populationsvarians**
- χ^2**-test för modellutvärdering**

Hypotestest

- Exempel 1. Genomsnittshastigheten på ditt bredband är sämre än leverantören utlovat.

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

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

- Exempel 2. En ny medicin påverkar blodtrycket.

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

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

- Exempel 3. En UI-förändring ökar andelen nöjda användare.

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

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

- Exempel 4. Andelen KD-väljare är under 4%-spärren.

$$\textbf{Nollhypotes}: \; H_0 : p \geq 0.04$$

$$\textbf{Alternativhypotes}: \; H_A : p < 0.04$$

Hypotestest

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

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

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

- **Ensidigt test**

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

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

  eller

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

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

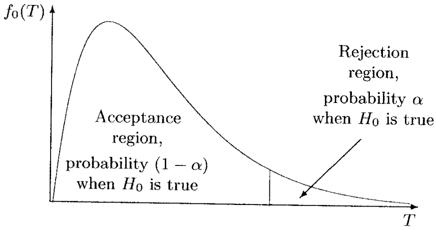
- Ensidiga test skrivs ibland så här (**det ger samma resultat**):

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

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

Steg vid hypotestest

1. Välj **teststatistika**, $T = T(X_1, \ldots, X_n)$.
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 $\boldsymbol{P}(T \in \mathcal{R} | H_0) = \alpha$.
4. Förkasta H_0 på **signifikansnivån** α om $T_{obs} \in \mathcal{R}$, där T_{obs} är det **observerade** värdet på T.



- H_0 gäller populationen. Då, H_0 är sann eller ej. Då, acceptera H_0 betyder inte att H_0 är sann med sannolikhet $1 - \alpha$. Det betyder att stickprovet stödjer H_0 tillräckligt för att inte förkasta den.
- Se Example 9.25 i Baron.

Hypotestest fel

- **Typ I fel**

$$\boldsymbol{\alpha} = \boldsymbol{P}(\text{Förkasta } H_0 | H_0 \text{ är sann})$$

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

- **Typ II fel**

$$\boldsymbol{P}(\text{Acceptera } H_0 | H_0 \text{ är falsk})$$

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

- **Styrka** (power): $\boldsymbol{P}(\text{Förkasta } H_0 | H_A \text{ är sann}) = 1$ - typ II fel.
- Rent formellt, förkastar man H_0 eller ej. Man aldrig accepterar den.

Steg vid hypotestest

- Allmänt: H_0 vs H_A

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

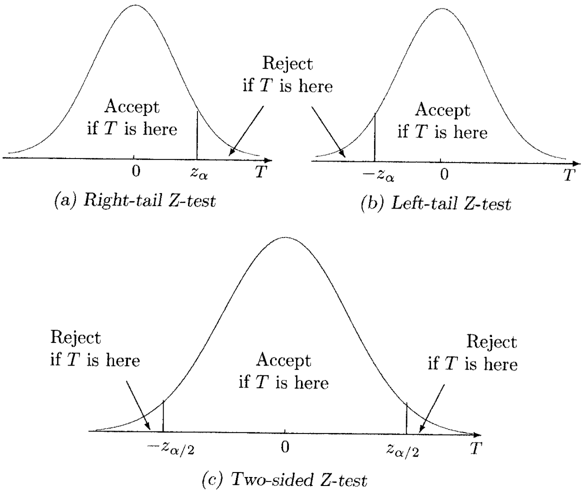
- Bernoulliexempel: $H_0 : p \leq 0.6$ vs $H_A : p > 0.6$.

  1. **Teststatistika**, $T = S = \sum_{i=1}^{n} X_i$.
  2. **Nollfördelningen** för T: $\sum_{i=1}^{n} X_i \sim Bin(n, 0.6)$.
  3. Låt $\alpha = 0.05$. Då, `qbinom(p=0.05,size=100,prob=0.6, lower.tail=F)` ger $\mathcal{R} = [68, 100]$ ungefär.
  4. Säg $T_{obs} = 70$. Då, $T_{obs} \in \mathcal{R}$ och $H_0 : p \leq 0.6$ förkastas på signifikansnivån 0.05.

Z-test

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

Z-test



(a) *Right-tail Z-test* (b) *Left-tail Z-test* (c) *Two-sided Z-test*

- Återbesök Example 9.25 i Baron.

Z-test för skillnad mellan väntevärdena av två populationer

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

$$H_0 : \mu_X = \mu_Y$$

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

- Exempel: Är genomsnittslönen lika stor i Stockholm och Göteborg?
- Låt $X_1, \ldots, X_n$ vara ett slumpmässigt stickprov från $N(\mu_X, \sigma_X^2)$.
- Låt $Y_1, \ldots, Y_m$ vara ett slumpmässigt stickprov från $N(\mu_Y, \sigma_Y^2)$.
- **Teststatistika**: $\bar{X} - \bar{Y}$. Samplingfördelning under H_0 ?
  - Linjärkombination av normalvariabler är normalfördelad, dvs $\bar{X} - \bar{Y}$ är normalfördelad.
  - $E(\bar{X} - \bar{Y}) = \mu_X - \mu_Y = 0$ under H_0.
  - $Var(\bar{X} - \bar{Y}) = Var(\bar{X}) + Var(\bar{Y}) = \sigma_X^2/n + \sigma_Y^2/m$.
  - Då,

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

- Se Example 9.26 i Baron.

T-test

- Z-test används när nollfördelningen är normalfördelad, dvs samplingfördelningen för teststatistikan om H_0 är sann är normalfördelad.
- Om σ^2 **inte** är känd utan skattas med s^2 blir inte

$$\frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

  längre normalfördelad utan t-**fördelad** med $n - 1$ frihetsgrader.

- Och z_α blir istället t_α och hämtas från Tabell A5 i Byron.

Koppling mellan hypotestest och konfidensintervall

Ett Z-test eller T-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 H_0 baserat på väldigt lite bevis.
- Idé: Presentera resultat för alla α.
- **P-värde** $=$ den lägsta signifikansnivån α där vi kan förkasta H_0.
- Alternativ definition: Sannolikheten att få en teststatistika som är lika extrem eller ännu mer extrem än T_{obs}.
- Exempel: Ensidigt Z-test. Då,

$$p\text{-värde} = \boldsymbol{P}(Z \geq Z_{obs})$$

- Se Example 9.38 i Baron.

χ^2-test för populationsvarians

- Väntevärdesriktig estimator av σ^2

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

- **Samplingfördelning** för s^2 om $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

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

- χ^2-fördelningen kan användas för konfidensintervall och hypotestest för σ^2. Se Sections 9.5.2 och 9.5.3 i Baron.
- $\chi^2_{1-\frac{\alpha}{2}}$ och $\chi^2_{\frac{\alpha}{2}}$ hämtas från Tabell A6 i Baron. Obs. båda behövs eftersom χ^2_ν-fördelningen är **inte symmetrisk**.
- χ^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

$$\mathbb{E}(X) = \nu \text{ och } Var(X) = 2\nu$$

- χ^2_ν-fördelningen är ett specialfall av gamma: $Gamma(\nu/2, 1/2)$.

χ^2 goodness of fit test

- Antag att populationen har följande diskreta fördelningen F_0, dvs $\boldsymbol{P}(X = 1) = p_1, \boldsymbol{P}(X = 2) = p_2, \ldots, \boldsymbol{P}(X = m) = p_m$.
- Om du har observerat n observationer så förväntar du dig $n \cdot p_k$ 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

- **χ^2-statistika**

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

- 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 χ^2-statistikan en χ^2-fördelning med $m - 1$ frihetsgrader, om $Exp(k) > 5$ för alla k.
- Vi kan även testa om data kommer från $F_0(\theta)$ där θ är en **okänd** parameter som skattas med en konsistent estimator. Då, χ^2-statistikan följer en χ^2-fördelning med $m - d - 1$ frihetsgrader där d är dimensionen av θ. **Hela fördelningsfamiljen** testas på en gång. Se Section 10.1.2 i Baron.
- **Kontinuerliga fördelningar** kan hanteras genom diskretisering (men se till att $Exp(k) > 5$ för alla k).

Översikt

- **Hypotesttest**
- **Z-test**
- **T-test**
- χ^2**-test för populationsvarians**
- χ^2**-test för modellutvärdering**