Statistik 1 för biologer, logopeder och psykologer

Föreläsningar, del 5

Innehåll

Innehåll

Inledande exempel

Exempel.

- Vi är intresserade av en variabel X om vilken vi kan anta att den är (approximativt) normalfördelad med standardavvikelsen $\sigma = 15$ i den givna populationen.
- På basen av tidigare erfarenheter är man benägen att tro att populationsmedelvärdet (dvs. väntevärdet μ) ligger kring 100.
- Det finns dock en misstanke om att det eventuellt skett en ökning i populationsmedelvärdet.
- För att undersöka saken drar man ett stickprov om $n = 25$ enheter och räknar medelvärdet på observationerna. Vi får ett stickprovsmedeltal på $\overline{x} = 118.7$, vilket är betydligt större än 100, men kan detta tas som ett bevis på att det faktiskt skett en förändring i populationen?

Inledande exempel (forts.)

Exempel.

Modell för populationen:

$$X \sim N(\mu, 15)$$

Stickprovsfördelningen för $\overline{X}$ är då $n = 25$:

$$\overline{X} \sim \; N(\mu, \frac{15}{\sqrt{25}}) \; .$$

Om vi nu utgår ifrån att populationsmedelvärdet verkligen är 100, dvs. att ingen förändring har skett, kan vi räkna ut sannolikheten för att t.ex. få ett stickprovsmedelvärde som är större än 115:

Vi har då $\mu = 100$ vilket betyder att $\overline{X} \sim N(100, 15/\sqrt{25})$. Sannolikheten för att medelvärdet av 25 observationer är större än 115 är i detta fall

$$P(\overline{X} > 115) = 1 - \Phi(\frac{115 - 100}{15/\sqrt{25}}) = 0.000000287 \; .$$

Inledande exempel (forts.)

Exempel.

Det observerade stickprovsmedelvärdet var 118.7, vilket ju är större än 115. Vi har då två alternativa förklaringar:

1. Något ytterst osannolikt har inträffat.
2. Vårt antagandde om att $\mu = 100$, dvs. att ingen förändring i populationsmedelvärdet har skett, är felaktigt.

Även om den första förklaringen är möjlig (också osannolika händelser kan ibland inträffa!) förefaller den senare förklaringen betydligt rimligare.

- Det antagande vars riktighet vi önskar pröva (testa) kallar vi nollhypotes
  - t.ex. "populationsmedelvärdet är 100".
- Ett alternativt antagande som vi har större benägenhet att tro på ifall nollhypotesen verkar orimlig kallar vi mothypotes
  - t.ex. "populationsmedelvärdet har ökat".

Översikt av stegen i hypotesprövning

1. Bestäm noll- och mothypotes.
2. Beräkna vad som kan förväntas om nollhypotesen vore sann.
3. Jämför det faktiska utfallet med vad som kan förväntas.
4. Om utfall och förväntan inte stämmer överens förkastar vi nollhypotesen och drar slutsatsen att mothypotesen är mera rimlig.

Hypotesprövning: Steg 1

Formulera en nollhypotes (betecknas H_0) och en mothypotes (betecknas H_1).

- Hypoteserna bör vara så formulerade att de täcker alla tänkbara möjligheter och är varandra uteslutande.
- Till H_0 väljes vanligen det som "gällt tidigare" eller det som man är beredd att hålla fast vid tills man blivit övertygad om att hypotesen är felaktig.
- Vid *jämförelser* är H_0 vanligen den att det inte finns några skillnader mellan de populationer som man jämför.
- Vid studier av *samband* formuleras H_0 vanligen som så att det inte finns något samband.

Hypotesprövning: Steg 2

Drag ett stickprov och beräkna värdet på någon lämplig teststatistika.

- Teststatistikan (kallas även testvariabel) är en sådan funktion av stickprovsobservationerna att man vet dess fördelning om H_0 är sann.
- Detta är nödvändigt för att vi skall kunna bedöma vad som är "sannolika" och "osannolika värden" på teststatistikan om H_0 vore sann.
- Ett exempel på en teststatistika är stickprovsmedelvärdet.

Hypotesprövning: Steg 3

Utgående från fördelningen för teststatistikan, dela in alla dess möjliga värden i sådana som är "sannolika" resp. "osannolika" om H_0 är sann.

- Ofta är teststatistikan sådan att både "väldigt små" och "väldigt stora" värden är osannolika om H_0 är sann. Ett test är då tvåsidigt.
- Om endast "väldigt små" eller "väldigt stora" värden är osannolika är ett test ensidigt.
- Formuleringen av mothypotesen H_1 bestämmer huruvida ett test är två- eller ensidigt.
- De värden i teststatistikans fördelning som kan anses vara "mycket osannolika" om H_0 är sann, utgör det sk. kritiska området.

Hypotesprövning: Steg 4

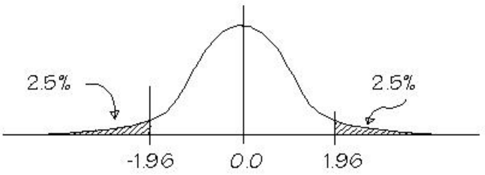
Om värdet på teststatistikan faller inom det kritiska området förkastas H_0.

- Det kritiska området brukar vanligen väljas så att sannolikheten för att få ett värde på teststatistikan som faller inom det kritiska området är bara 0.05 eller 0.01 om H_0 är sann.
- Sannolikheten för att observera ett värde på teststatistikan som faller inom det kritiska området givet att H_0 är sann kallas testets signifikansnivå och betecknas α.
    - Man talar ofta om t.ex. ”test på 5%-nivån” eller ”test på 1%-nivån”.
- Ett värde på teststatistikan som faller inom det kritiska området sägs vara signifikant.

OBS! Att H_0 inte förkastas bevisar ingalunda att nollhypotesen är sann.

Hypotesprövning

Kritiskt område för ett tvåsidigt test på 5%-nivån ($\alpha = 0.05$) då teststatistikan följer en normalfördeling.



Typ I och typ II -fel och testets styrka

- Testets signifikansnivå α kan även tolkas som sannolikheten att förkasta H_0 då den är sann, dvs. sannolikheten för ett typ I -fel.
- I hypotesprövning kan man också begå ett annat slags fel, nämligen att låta bli att förkasta H_0 då H_1 är sann. Sannolikheten för det sk. typ II -felet brukar betecknas β.
- De två felen hänger ihop så att om man minskar risken för den ena kommer risken för den andra att öka. Genom att öka på stickprovsstorleken kan man dock minska båda riskerna samtidigt.
- Testets styrka är dess förmåga att förkasta en felaktig nollhypotes. Då H_1 är sann definieras styrkan som $1 - \beta$.

Typ I och typ II -fel och testets styrka

Tabellen nedan sammanfattar de olika besluten samt deras konsekvenser vid hypotesprövning:

Verklighet	Beslut	
	H_0 **förkastas**	H_0 **förkastas ej**
H_0 **sann**	Typ I -fel (α-fel)	Korrekt
H_1 **sann**	Korrekt	Typ II -fel (β-fel)

p-värden

- Vi påminner oss om att ju mera extremt ("väldigt stort" och/eller "väldigt litet") värde på teststatistikan vi observerar, desto mindre stöd får H_0.
- p-värdet är sannolikheten för att erhålla ett värde på teststatistikan som är minst lika extremt som det som man faktiskt observerat givet att H_0 är sann.
- T.ex är då
  - ett p-värde som är ≤ 0.05 är signifikant på 5%-nivån
  - ett p-värde som är ≤ 0.01 är signifikant på 1%-nivån.

Vidare om p-värden och signifikans

- Ett statistiskt signifikant resultat behöver inte nödvändigtvis ha någon praktisk relevans.
  - T.ex. då man jämför värdet på en variabel hos två mycket stora grupper kan man lätt upptäcka en statistiskt signifikant skillnad som dock är så liten att den i praktiken inte har någon relevans.
- Då uppdelningen i "signifikanta" och "icke-signifikanta" resultat är något konstgjord, är det i många fall mera meningsfullt att endast rapportera det erhållna p-värdet och sedan verbalt tolka huruvida resultatet talar emot H_0.

Om tolkning av p-värdet

- En vanlig missuppfattning är att tro att p-värdet är anger sannolikheten för att H_0 är sann. **Detta stämmer alltså inte.**
- I stället anger p-värdet hur trovärdigt det observerade datamaterialet är *ifall* H_0 vore sann. Föga trovärdiga data (litet p-värde) talar då alltså emot H_0.
- Mer om p-värden, se
  - `http://en.wikipedia.org/wiki/P_value`

  samt speciellt den i Wikipedia-artikeln ovan citerade artikeln
  - `http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=11159626`.

Test av väntevärde då σ är känt

- Vi har ett stickprov från en population vars väntevärde μ är okänt medan standardavvikelsen σ antas vara *känd*, och vi önskar testa hypotesen
$$H_0 : \mu = \mu_0 \ .$$
- En lämplig testvariabel är då
$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1) \ .$$
- I ett *tvåsidigt* test lägger vi upp mothypotesen
$$H_1 : \mu \neq \mu_0 \ .$$
- I ett test på signifikansnivå α förkastar vi H_0 om $z < -z_{\alpha/2}$ eller $z > z_{\alpha/2}$, där några vanligt använda värden för $z_{\alpha/2}$ är:

$z_{\alpha/2}$	1.96	2.54	3.29
α	0.05	0.01	0.001

Test av väntevärde då σ är känt

- Har vi i stället ett *ensidigt* test formuleras mothypotesen antingen som
$$H_1 : \mu < \mu_0 \ .$$
eller som
$$H_1 : \mu > \mu_0 \ .$$
- Om $H_1 : \mu < \mu_0$ förkastas H_0 ifall $z < -z_\alpha$ och om $H_1 : \mu > \mu_0$ förkastas H_0 ifall $z > z_\alpha$. Några vanligt använda värden för z_α är:

z_α	1.64	2.33	3.09
α	0.05	0.01	0.001

Test av väntevärde då σ är okänt

- Vi har ett stickprov från en *normalfördelad* population vars väntevärde μ och standardavvikelse σ är *okända*, och vi önskar testa hypotesen
$$H_0 : \mu = \mu_0 \ .$$
- En lämplig testvariabel är då
$$T = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} \ ,$$
vilken följer en t-fördelning med frihetsgraderna $f = n - 1$.
- De två- och ensidiga mothypoteserna lägger vi upp som tidigare.
- I ett tvåsidigt test på signifikansnivå α förkastas H_0 om $t < -t_{\alpha/2}$ eller $t > t_{\alpha/2}$, där $t_{\alpha/2}$ är ett sådant värde av t-fördelningen att
$$P(T < -t_{\alpha/2}) + P(T > t_{\alpha/2}) = \alpha \ .$$

Test av väntevärde då σ är okänt

- I ett ensidigt test med mothypotesen $H_1 : \mu < \mu_0$ och signifikansnivån α förkastas H_0 om $t < -t_\alpha$.
- I ett ensidigt test med mothypotesen $H_1 : \mu > \mu_0$ och signifikansnivån α förkastas H_0 om $t > t_\alpha$.
- t_α är ett sådant värde av t-fördelningen att

$$P(T > t_\alpha) = \alpha \ .$$

- Det ovan beskrivna testet kallas ofta för t-test, eftersom testvariabeln följer en t-fördelning.

Test av väntevärde då σ är okänt – exempel

Exempel.

- Vi antar att variablen X i en population är normalfördelad med μ och σ okända. Vi drar ett stickprov om $n = 60$ enheter och räknar stickprovsmedelvärdet och -standardavvikelsen, vilka blir $\bar{x} = 16.51$ resp. $s = 1.23$.
- Låt oss nu pröva hypotesen

$$H_0 : \mu = 17.0 \quad \text{mot}$$
$$H_1 : \mu \neq 17.0 \,.$$

- Vi har då ett tvåsidigt test och bör alltså förkasta H_0 om $\bar{x}$ är antingen tillräckligt stort eller tillräckligt litet.

Test av väntevärde då σ är okänt – exempel (forts.)

Exempel.

- Vi räknar nu värdet på testvariabeln
$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{16.51 - 17.0}{1.23/\sqrt{60}} = -3.09 \ .$$
- Det erhållna värdet jämför vi med det kritika värdet för signifikansnivån $\alpha = 0.05$ i en t-fördelning med frihetsgraderna $f = 60 - 1$. T.ex. med hjälp av en tabell finner vi att det kritiska värdet i detta fall är $t_{\alpha/2} = 2.00$
- Eftersom teststatistikans värde $t = -3.09 < -t_{\alpha/2} = -2.00$ förkastar vi H_0.
- Vi kunde alternativt ha baserat beslutet på att det tvåsidiga p-värdet 0.003 är mindre än 0.05.

Jämförelse av två väntevärden då σ är okänt

- Vi har två *normalfördelade* populationer och drar ett stickprov från vardera. Låt storleken på stickproven vara n_1 respektive n_2.
- Vi antar att populationernas standardavikelser σ_1 resp. σ_2 är okända men har orsak att tro att de är (nästan) lika, dvs. $\sigma_1 = \sigma_2 = \sigma$.
- Vi önskar testa hypotesen

$$H_0 : \mu_1 = \mu_2 \ ,$$

vilken även kan formuleras som $H_0 : \mu_1 - \mu_2 = 0$.

Jämförelse av två väntevärden då σ är okänt

- En lämplig testvariabel är då

$$T = \frac{\overline{X}_1 - \overline{X}_2}{s_p / \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

där s_p är den "poolade" standardavvikelsen. T följer då en t-fördelning med frihetsgraderna $f = n_1 + n_2 - 2$.

- Mothypotesen är någon av följande:
  - $H_1 : \mu_1 \neq \mu_2$
  - $H_1 : \mu_1 < \mu_2$
  - $H_1 : \mu_1 > \mu_2$.
- H_0 förkastas enligt samma principer som i det nyss introcuerade t-testet (även detta är ett t-test).

Jämförelse av proportioner

- Vi har två populationer. I vardera populationen har en okänd andel av enheterna en viss egenskap. Vi betecknar andelarna p_1 och p_2.
- Vi drar var sitt stickprov av storlek n_1 resp. n_2 från de två populationerna. Låt X_1 beteckna antalet enheter i det första stickprovet med den sökta egenskapen och låt X_2 beteckna motsvarande antal i det andra stickprovet.
- Andelen enheter med egenskapen i vardera stickprov ges då av $\frac{\overline{X}_1}{n_1}$ resp. $\frac{\overline{X}_2}{n_2}$.
- Vi vill nu undersöka om andelarna i de två populationerna är lika. Detta motsvaras av nollhypotesen

$$H_0 : p_1 = p_2 ,$$

vilken även kan formuleras som $H_0 : p_1 - p_2 = 0$.

Jämförelse av proportioner

- En lämplig testvariabel är då

$$Z = \frac{\frac{X_1}{n_1} - \frac{X_2}{n_2}}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}},$$

där $\hat{p} = \frac{X_1+X_2}{n_1+n_2}$. Variabeln Z följer en $N(0,1)$-fördelning.

- Mothypotesen är någon av följande:
  - $H_1 : p_1 \neq p_2$
  - $H_1 : p_1 < p_2$
  - $H_1 : p_1 > p_2$.
- Eftersom testvariabeln följer en standardiserad normalfördelning, förkastas H_0 enligt samma principer som t.ex. vid test av väntevärde då σ är känt.

Jämförelse av proportioner – exempel

Exempel.

- I ett förberedande skede av en genetisk studie vill man ta reda på om andelen brunögda invånare i två städer A och B är lika. De okända proportionerna betecknas p_A resp. p_B.
- Ett stickprov dras ur populationerna i vardera städer. I stad A omfattar stickprovet $n_A = 150$ individer, av vilka antalet brunögda är $x_A = 103$. Motsvarande tal för stad B är $n_B = 120$ och $x_B = 77$
- Vi lägger upp hypoteserna

$$
\begin{aligned}
H_0 &: p_A = p_B \\
H_1 &: p_A \neq p_B \ .
\end{aligned}
$$

Jämförelse av proportioner – exempel (forts.)

Exempel.

- Testvariabeln räknar vi enligt den angivna formeln

$$z = \frac{\frac{103}{150} - \frac{77}{120}}{\sqrt{\frac{2}{3} \cdot \frac{1}{3} \cdot \left(\frac{1}{150} + \frac{1}{120}\right)}} \approx 0.78$$

- Testar vi hypotesen på signifikansnivån $\alpha = 0.05$ får vi att $z = 0.78 < z_\alpha = 1.96$, alltså låter vi bli att förkasta H_0.

Test av regressionskoefficient

- Vi har i början av kursen bekantat oss med regressionsmodeller där man är intresserad av ett funktionellt samband mellan en beroende variabel och en eller flera förklarande variabler.
- Låt oss nu anta att två variabler X och Y i en population har följande linjära samband

$$Y = \alpha + \beta \cdot X + (\text{slumpmässigt fel}) \, .$$

- α och β är här okända parametrar vars värden skattas från ett stickprov. I praktiken görs detta oftast med hjälp av ett statistiskt programpaket.

Test av regressionskoefficient

- I utskriften av en dylik regressionsanalys brukar det utöver det skattade värdet på regressionskoefficienten även anges ett p-värde för koefficienten.
- Detta p-värde hör ihop med ett test av hypoteserna

$$\begin{aligned} H_0 &: \beta = 0 \\ H_1 &: \beta \neq 0 \,, \end{aligned}$$

  där H_0 innebär att det inte finns något samband mellan variablerna X och Y medan H_1 innebär att det finns ett samband.
- Testet är baserat på en t-fördelning och har den bekanta tolkningen av att ju mindre p-värdet är desto mer talar det emot H_0.

Förteckning över vanliga test

En förteckning över vanliga test hittas på sidan

`http://en.wikipedia.org/wiki/Hypothesis_testing#Common_test_statistics`

Icke-parametriska test

- Då vi inte testar värdet på en parameter utan är är t.ex. intresserade av formen på en fördelning talar vi om ett icke-parametriskt test.
- Icke-parametriska test används ofta då man inte kan göra antaganden om normalfördeling eller då data är mätta på nominal- eller ordinalskala.
- Ett par vanliga icke-parametriska test är
  - Mann-Whitney-Wilcoxon testet
    `http://en.wikipedia.org/wiki/Mann-Whitney-Wilcoxon`
  - Kolmogorom-Smirnov testet
    `http://en.wikipedia.org/wiki/Kolmogorov-Smirnov_test`
- Ett av de mest använda icke-parametriska testen är det sk. χ^2-testet, vilket det finns flera olika varianter av.

Test av fördelning med χ^2-test

- För att testa hur väl frekvensfördelningen på ett klassificerat datamaterial stämmer överens med en förväntad fördelning kan man använda ett χ^2-test.
- Hypoteserna i ett dylikt test läggs upp som följande:
  - H_0 : Den förväntade och observerade fördelningen är lika.
  - H_1 : Den förväntade och observerade fördelningen är olika.
- Värdet på testvariabeln beräknas enligt formeln

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \; ,$$

där O_i = observerad frekvens och E_i = förväntad frekvens under antagande att H_0 är sann. Summeringen sker över alla klasser i fördelningen.

Test av fördelning med χ^2-test

- Då H_0 är sann, följer testvariabeln en χ^2-fördelning med frihetsgraderna $f = k - r - 1$, där k är antalet klasser och r är antalet parmametrar som har skattats från datamaterialet.
- I det allra enklaste fallet jämför vi den observerade fördelningen med en likformig fördelning, vilket inte kräver skattning av parametrar. Frihetsgraderna blir då $f = k - 1$.
- H_0 förkastas om testvariabelns värde χ^2 överskrider det signifikansnivån α motsvarande kritiska värdet χ^2_α i en χ^2-fördelning med frihetsgraderna f.

Test av fördelning med χ^2-test – exempel

Exempel.

- Ledningen för ett företag misstänker att en del anställda har tagit för vana att förlänga veckoslutet genom att sjukanmäla sig för fredag och/eller måndag.
- Saken följdes upp under fyra veckors tid.
- Ifall ledingens misstanke är obefogad, skulle vi förvänta oss att sjukfrånvarona fördelar sig jämnt över alla arbetsdagar.
- Resultatet blev följande:

Veckodag	må	ti	on	to	fre
Observerade frånvaron	49	35	32	39	45
Förväntade frånvaron	40	40	40	40	40

Test av fördelning med χ^2-test – exempel

Exempel.

- Vi lägger upp följande hypoteser:

$$\begin{aligned} H_0 &: \text{ Frånvarona fördelar sig jämnt över alla veckodagar.} \\ H_1 &: \text{ Frånvarona fördelar sig inte jämnt över alla veckodagar.} \end{aligned}$$

- Enligt den tidigare angivna formeln blir värdet på testvariabeln

$$\chi^2 = \frac{(49-40)^2}{40} + \frac{(35-40)^2}{40} + \cdots + \frac{(45-40)^2}{40} = 4.9 \ .$$

- Eftersom den förväntade fördelningen är likformig behöver inga parametrar skattas. Vi jämför således testvariabelns värde med en χ^2-fördelning med frihetsgraderna $f = 5 - 1 = 4$.
- Väljer vi signifikansnivån $\alpha = 0.05$ får vi det motsvarande kritska värdet 9.488 t.ex ur en tabell. Eftersom testvariabelns värde inte överskrider det kritiska värdet förkastar vi inte H_0.