

TMS136

Föreläsning 11

Andra intervallskattningar

- Vi har sett att vi givet ett stickprov och under vissa antaganden kan göra intervallskattningar för väntevärden
- Man kan även göra intervallskattningar för varianser i normalfördelningar och för proportioner i binomialfördelningar

Konfidensintervall för varians i normalfördelning

- Vi vet att vi givet ett stickprov $X_1, \dots, X_n$ kan skatta populationsvariansen σ^2 med stickprovsvariansen

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

Konfidensintervall för varians i normalfördelning

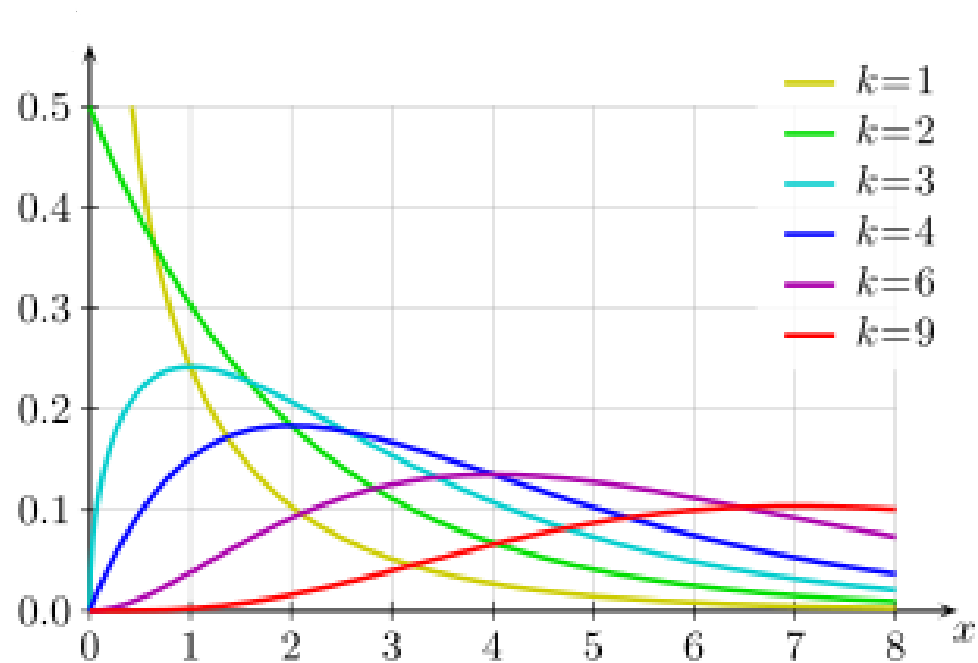
- Om stickprovet kommer från en normalfördelning gäller att

$$\frac{(n - 1)S^2}{\sigma^2}$$

har χ^2 -fördelning ("chi-två") med $n - 1$ frihetsgrader

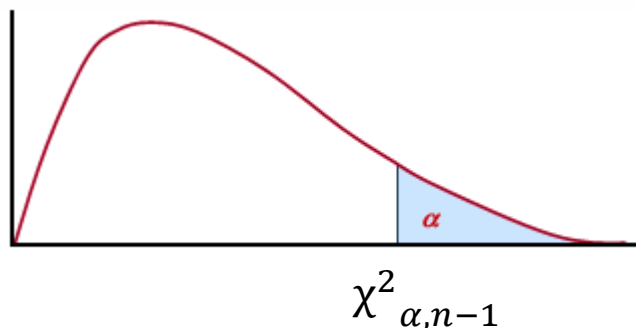
χ^2 -fördelning

- Täthetsfunktionen för χ^2 -fördelningen med 1, 2, 3, 4, 6 och 9 frihetsgrader



χ^2 -fördelning

- Vi ser att χ^2 -fördelade s.v. är positiva och att tätheterna är skeva (asymmetriska)
- Precis som för normalfördelning och t -fördelning finns inga slutna uttryck för fördelningsfunktionen så man måste slå i tabeller eller använda numerik om man vill bestämma sannolikheter för χ^2 -fördelade s.v.
- Det tal som avgränsar arean under täthetsfunktionens graf så att $\alpha \cdot 100\%$ av arean finns ovanför talet kallar vi $\chi^2_{\alpha, n-1}$



Konfidensintervall för varians i normalfördelning

- Så om vi vill göra ett $(1 - \alpha) \cdot 100\%$ -igt konfidensintervall för variansen i en normalfördelning ska vi alltså "klämma in"

$$\frac{(n - 1)S^2}{\sigma^2}$$

mellan de tal $\chi^2_{1-\alpha/2, n-1}$ och $\chi^2_{\alpha/2, n-1}$ sådana att

$$P\left(\chi^2_{1-\alpha/2, n-1} \leq \frac{(n - 1)S^2}{\sigma^2} \leq \chi^2_{\alpha/2, n-1}\right) = 1 - \alpha$$

Konfidensintervall för varians i normalfördelning

- Möblerar vi om i parentesen får vi

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2, n-1}}\right) = 1 - \alpha$$

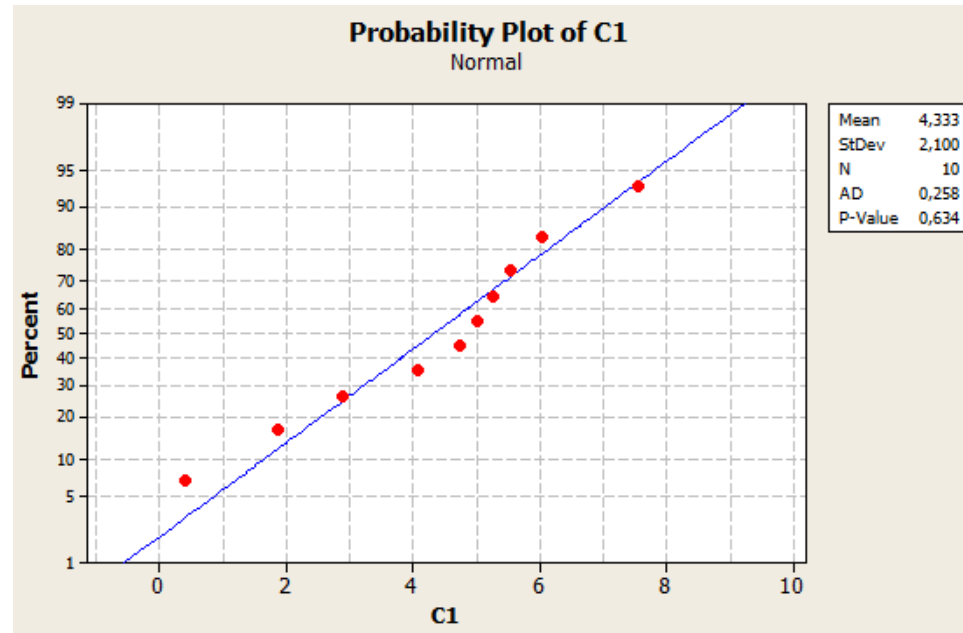
- Om vi observerar $x_1, \dots, x_n$ observerar vi ju också s^2 och konfidensintervallet

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}} \right]$$

Exempel

- Vi observerar stickprovet

4,74
2,88
0,41
6,02
1,88
5,55
5,25
4,07
5,00
7,54



- Stickprovet passerar ett normalfördelningstest

Exempel

- Vi vill skapa ett 95%-igt konfidensintervall för variansen i den normalfördelning som stickprovet antas komma ifrån
- Vi har att $s^2 \approx 4.41$ och $n = 10$ ger vid koll i appendix att $\chi^2_{0.025,9} = 19.02$ och $\chi^2_{0.975,9} = 2.70$
- Det följer att konfidensintervallet blir

$$\left[\frac{9 \cdot 4.41}{19.02}, \frac{9 \cdot 4.41}{2.70} \right] \approx [2.09, 14.7]$$

Konfidensintervall för σ

- Om man istället vill ha konfidensintervall för standardavvikelsen i en normalfördelning gör man ett för variansen och ”drar roten ur”
- I exemplet ovan får vi alltså ett 95%-igt konfidensintervall för σ som är

$$[\sqrt{2.09}, \sqrt{14.7}] \approx [1.45, 3.83]$$

Ensidiga konfidsensgränser för σ^2 och σ

- Om vi istället vill göra ensidiga $(1 - \alpha) \cdot 100\%$ -iga konfidsensgränser får vi dessa genom att "lägga hela α på ena sidan"
- Alltså, en undre respektive övre $(1 - \alpha) \cdot 100\%$ -ig konfidsensgräns för σ^2 ges av

$$\frac{(n-1)s^2}{\chi^2_{\alpha, n-1}} \text{ respektive } \frac{(n-1)s^2}{\chi^2_{1-\alpha, n-1}}$$

- Motsvarande för σ fås genom att "dra roten ur"

Exempel

- Med samma data som i det tvåsidiga exemplet får vi en övre 95% konfidsgräns för σ^2 genom att leta upp $\chi^2_{0.95,9} = 3.33$ i appendix och ”pussla ihop” enligt

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha, n-1}} = \frac{9 \cdot 4.41}{3.33} \approx 11.9$$

Konfidensintervall för proportioner

- Vi vet att om parameteren n i en binomialfördelning är stor så kan binomialfördelningen approximeras med normalfördelning (CGS)
- Detta betyder att om vi har stora stickprov så kan vi med hjälp av normalfördelning "klämma in" en populationsproportion motsvarande parameteren p i en binomialfördelning i ett konfidensintervall

Konfidensintervall för proportioner

- Alltså: Om X är binomialfördelad med parametrar n och p , där n är stort, och vi skattar p med $\hat{p} = X/n$ så är

$$\frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

approximativt standard normalfördelad, vilket ger att

$$P\left(-z_{\alpha/2} \leq \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{\alpha/2}\right) \approx 1 - \alpha$$

Konfidensintervall för proportioner

- Möblerar vi om får vi att

$$P\left(\hat{p} - z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \leq p \leq \hat{p} + z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}\right) \approx 1 - \alpha$$

För att kunna skapa konfidensintervallet från observerade data måste p ersättas med $\hat{p}$ under rottecknet

- Så det approximativt $(1 - \alpha) \cdot 100\%$ -iga konfidensintervallet fås ur

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

Exempel

- Från ett stickprov av storlek 200 observerar vi att 82 personer är för trängselskatt i Göteborg
- Vi vill göra ett (approximativt) 95%-igt konfidensintervall för hur stor andel av Göteborgs befolkning som är för trängselskatt

Exempel

- Vi observerar alltså att $\hat{p} = \frac{82}{200} = 0.41$
- Vi vet att $z_{0.025} = 1.96$ så det sökta konfidensintervallet blir

$$0.41 \pm 1.96 \sqrt{\frac{0.41 \cdot 0.59}{200}}$$

alltså $[0.34, 0.48]$

Hur stort stickprov behövs för en viss intervallbredd?

- Om vi vill att konfidensintervallet för proportionen ska ha bredden $2E$ kan vi få detta om vi väljer stickprovsstorleken (verifiera)

$$n = \left(\frac{Z_{\alpha/2}}{E} \right)^2 p(1 - p)$$

- Detta förutsätter dock att vi har en uppfattning om vad p är. Det kanske vi har från en pilotstudie...

Hur stort stickprov behövs för en viss intervallbredd?

- Om vi inte har någon uppfattning om vad p är vet vi alla fall att $p(1 - p) \leq 0.25$
- Så "worst case" är att

$$n = 0.25 \cdot \left(\frac{Z_{\alpha/2}}{E} \right)^2$$

Exempel

- Vi vill göra om undersökningen om trängselskatten så att vi får ett 95%-igt konfidensintervall som är 1 procentenhet brett
- Om vi använder den förra studien som "riktmärke" får vi att vi ska använda ett stickprov av storlek (avrundat upp till närmsta hundratal)

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{1.96}{0.005}\right)^2 \cdot 0.41 \cdot 0.59 \approx 37200$$

- Utan "riktmärket" skulle vi få

$$n = 0.25 \cdot \left(\frac{z_{\alpha/2}}{E}\right)^2 = 0.25 \cdot \left(\frac{1.96}{0.005}\right)^2 \approx 38500$$

Hypotestest

- I exempelvis en tillverkningprocess av 33cl läskburkar kan vi ha anledning att tro att burkarna fylls på lite för mycket
- Vi vill testa sanningshalten i detta med hjälp av stickprovsdata
- Vi sätter upp en *nollhypotes* (*null hypothesis*), H_0 , och en *mothypotes* (*alternative hypothesis*), H_1

Hypotesprövning

- I fallet med läskburkarna är det väntevärdet, μ , för hur mycket läsk som finns i burkarna, vi ska "testa"
- Hypoteserna blir

$$H_0: \mu = 33$$

$$H_1: \mu > 33$$

Mothypotesen är "det vi tror"

Hypotesprövning

- För stickprovet $X_1, \dots, X_n$ som antas komma från en normalfördelning vet vi att kvoten

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

är t -fördelad med $n - 1$ frihetsgrader

Hypotesprövning

- Så för att utföra hypotesprövningen ska vi jämföra det observerade värdet på kvoten med en t -kvantil, $t_{\alpha, n-1}$, alltså det tal som har $\alpha \cdot 100\%$ av arean under täthetsfunktionen till höger om sig
- Det är vanligt att man väljer $\alpha = 0.05$
- Nedan ska vi se vad ett specifikt val av α har för implikationer

Hypotesprövning

- Om det observerade värdet för kvoten är större än $t_{\alpha, n-1}$ *förkastar* vi nollhypotesen. Vi drar alltså slutsatsen att burkarna fylls för mycket
- Om det observerade värdet på kvoten inte är större än $t_{\alpha, n-1}$ kan vi inte förkasta nollhypotesen och har därmed inte heller något fog för att påstå att burkarna fylls för mycket

Hypotesprövning

- Vilken t -kvantil ska man välja?
- Vid hypotesprövning finns följande möjligheter

	H_0 är sann	H_0 är falsk
Inte förkasta H_0	Korrekt	Typ II-fel
Förkasta H_0	Typ I-fel	Korrekt

- Sannolikheten att göra Typ I-fel kallas *signifikansnivå*, betecknas α , och bestämmer vilken t -kvantil vi ska jämföra kvoten med
- Sannolikheten att gör Typ II-fel betecknas β

Hypotesprövning

- Antag att vi bestämmer oss för $\alpha = 0.05$ och att vi bland 30 slumpmässigt valda läsburkar observerar $\bar{x} = 33.3$ och $s = 0.36$
- Det observerade värdet på kvoten blir då

Stickprovsmedelvärdet Nollhypotesens värde på μ

$$\frac{33.3 - 33}{0.36 / \sqrt{30}} \approx 4.56$$

Stickprovsstandardavvikelsen Stickprovsstorleken

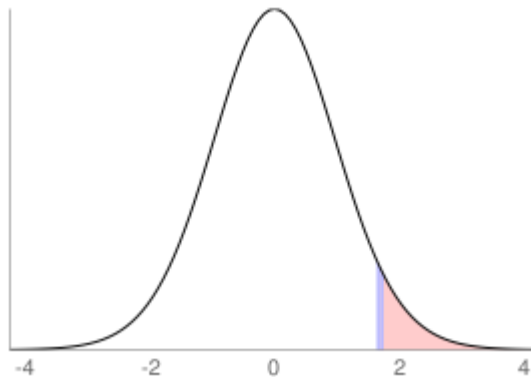
The diagram illustrates the calculation of a test statistic. It features a central equation: $\frac{33.3 - 33}{0.36 / \sqrt{30}} \approx 4.56$. Four red labels with blue arrows point to specific parts of the equation: 'Stickprovsmedelvärdet' points to 33.3, 'Nollhypotesens värde på μ ' points to 33, 'Stickprovsstandardavvikelsen' points to 0.36, and 'Stickprovsstorleken' points to 30.

Hypotesprövning

- Värdet på kvoten (4.56) ska alltså jämföras med t -kvantilen för 29 frihetsgrader och $\alpha = 0.05$
- Vi finner $t_{\alpha, n-1} = t_{0.05, 29} = 1.699$ i appendix
- Vi kan alltså förkasta nollhypotesen och gott påstå att burkarna fylls för mycket

Förkastelseområde

- I fallet med läskburkarna har vi *förkastelseområdet* (*critical region, region of rejection*) som är skuggat i bilden



- Det skuggade området är alltså (i det här fallet) sannolikheten att en t_{29} -fördelad slumpvariabel tar ett värde större än 1.699

p-värde

- I hypotesprövningssammanhang förekommer ofta uttrycket *p-värde*
- p-värdet säger hur sannolikt det är att vårt observerade stickprov ser ut som det gör (eller "värre") om nollhypotesen är sann
- Vi förkastar nollhypotesen om p-värdet är mindre än α

p-värde

- I fallet med läskburkarna blir p-värdet (räkna i exempelvis Excel; 1-T.DIST(4,56;29;TRUE))

$$P(t_{29} > 4.56) = 0.000043$$

- Det är alltså högst osannolikt att läskburkarna kommer från en fördelning med väntevärde 33cl

Olika hypoteser

- Vi kan pröva olika hypoteser kring en populationsparameter θ
- Ensidiga

$$H_0: \theta = \theta_0$$

$$H_1: \theta > \theta_0$$

eller

$$H_0: \theta = \theta_0$$

$$H_1: \theta < \theta_0$$

Olika hypoteser

- Tvåsidiga

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_0$$

- För läskburkarna kanske vi istället vill pröva

$$H_0: \mu = 33$$

$$H_1: \mu \neq 33$$

Fler läskburkar och tvåsidigt

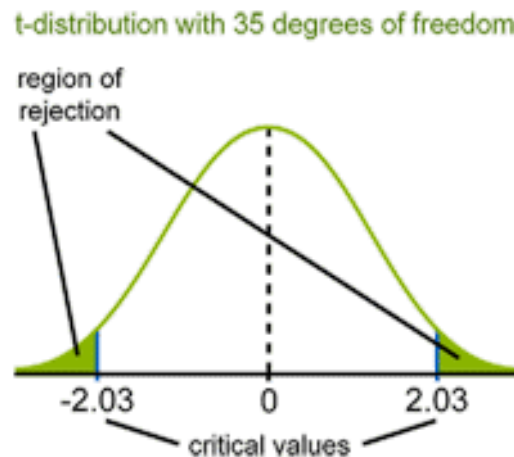
- Om vi fortsatt använder $\alpha = 0.05$ men tar ut ett stickprov om 36 burkar, får $\bar{x} = 33.2$ och $s = 0.38$, ska vi förkasta nollhypotesen om

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} > t_{35,0.025} \quad \text{eller} \quad \frac{\bar{x} - \mu}{s/\sqrt{n}} < -t_{35,0.025}$$

- $t_{35,0.025} \approx 2.03$ (appendix) och $\frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{33.2 - 33}{0.38/\sqrt{36}} \approx 3.16$ så vi förkastar nollhypotesen

Förkastelseområde

- När vi testar tvåsidigt med signifikansnivå $\alpha = 0.05$ är förkastelseområdet de 2.5% av sannolikheten längst till vänster och längst till höger



Hypotesprövning vs. konfidensintervall

- I fallet med läskburkarna skulle vi få det 95%-konfidensintervallet

$$(33.2 - 2.03 \cdot 0.38/\sqrt{36}, 33.2 + 2.03 \cdot 0.38/\sqrt{36})$$

$$\approx (33.07, 33.33)$$

- Nollhypotesens $\mu = 33$ finns alltså inte med i intervallet
- När det gäller väntevärden hos normalfördelningar kan man lika gärna göra (tvåsidiga) konfidensintervall som (tvåsidig) hypotesprövning

Tänkbara parametrar för hypotestest

- Väntevärde, μ , hos normalfördelning med känd varians testas mha den standard normalfördelade kvoten

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

och det observerade värdet utvärderas mot standard normalkvantil

Tänkbara parametrar för hypotestest

- Väntevärde, μ , hos normalfördelning med okänd varians testas som vi sett mha den t_{n-1} -fördelade kvoten

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

och det observerade värdet utvärderas mot t_{n-1} -kvantil

Tänkbara parametrar för hypotestest

- Proportionen, p , hos binomialfördelning testas mha den approximativt standard normalfördelade kvoten

$$\frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

och det observerade värdet utvärderas mot standard normalkvantil

Hypotest för proportion

- Vi tror att andelen läskburkar som är överfyllda är större än 0.4
- Vi ställer upp hypoteserna

$$H_0: p = 0.4$$

$$H_1: p > 0.4$$

- Vi tar ett stickprov om 28 burkar och upptäcker att 15 burkar är överfyllda

Hypotest för proportion

- Det observerade värdet för kvoten blir alltså

$$\frac{15/28 - 0.4}{\sqrt{0.4 \cdot 0.6/28}} \approx 1.466$$

- Om vi arbetar med $\alpha = 0.05$ är normalkvantilen $z_{0.05} = 1.64$ så vi kan inte förkasta nollhypotesen

Reflektion

- Att man inte kan förkasta nollhypotesen betyder inte nödvändigtvis att nollhypotesen är sann, bara att man inte har "nog bevis" att förkasta den
- Om man bara har tillräckligt stora stickprov kan man få signifikans för vilken som helst avvikelse från nollhypotesen
- Frågan är vad som är kliniskt relevant/har praktiska eller funktionella implikationer