Konfidensintervall

Om vi är intresserade av en viss proportion p (så som vilket valt som vi det) kan vi beräkna en sträcka av vilken stickprovsmedelvärdet $\bar{x}$ $\bar{x}$

1. om vi $n \geq 30$ så kommer $\bar{x} \to \bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$
2. om vi $np \geq 5$ så kommer $\bar{x} \to \bar{x} \sim N(p, \frac{\sigma}{\sqrt{n}})$
3. om vi $np \geq 5$ och $n(1-p) \geq 5$ kommer $\bar{x} \to \bar{x} \sim N(p, \frac{\sigma}{\sqrt{n}})$

Ok, om vi vet hur en population $X \sim N(\mu, \sigma)$ är kan vi bara räkna ut hur det ser ut för $p = 50$ gånger (vilket om vi vet vet det).

Men om vi vill se stickprov med mindre antal som vi inte vet hur det ser ut $\bar{x} \sim N$ gånger.

Vi vet att $\bar{x} \sim N(p, \frac{p}{\sqrt{n}})$

Genom $\bar{x} = 1.96$ gånger vi det minst eller mer p. Det vi kopar: $\bar{x} \approx p$

Vi kan hitta med konfidensintervall som det vi vet konfidens

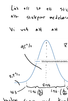
Resultat: Vi är med 95% sannolikhet säkra på att populationen kommer parameter p ligger i intervallet $[\bar{x} \pm 1.96 \cdot \frac{p}{\sqrt{n}}, \bar{x} \pm 1.96 \cdot \frac{p}{\sqrt{n}}]$

$\rightarrow$ $[463, 469.67]$

detta är vårt konfidensintervall

Antag att vi vill ha ett konfidensintervall som vi är dem säkrare:

Antag att vi vill ha en konfidens på 99%



Vi är 99% säkra på att populationen kommer p ligger i intervallet $[\bar{x} \pm 2.58 \cdot \frac{p}{\sqrt{n}}]$

> **Konfidensintervall** är ett begrepp som anger osäkerhet i statistiskt generaliserande. Det är ett intervallskattning som ger ett intervall som med en viss sannolikhet innehåller populationsparametern θ (läs: theta).
>
> Ett 95%-konfidensintervall ger en uppfattning om det intervall som kommer att innehålla den sanna populationsparametern θ i 95 procent av alla fall.
>
> $$\bar{x} \pm 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$
>
> $\sim 95\%$

om vi $n \geq 30$ kan vi se en normalfördelad population

$n \geq 30$ och $np \geq 5$ $\rightarrow$ $[\bar{x} - 1.96(\frac{\sigma}{\sqrt{n}}), \bar{x} + 1.96(\frac{\sigma}{\sqrt{n}})]$ $\rightarrow$ hur 95% population ligger

$\cdot$ $[\bar{x} - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \cdot \frac{\sigma}{\sqrt{n}}]$

$n < 30$ och $np \geq 5$ $\rightarrow$ $[50 - 1.96(\frac{\sigma}{\sqrt{n}}), 50 + 1.96(\frac{\sigma}{\sqrt{n}})]$

$\cdot$ $[50 - 1.96 \cdot \frac{\sigma}{\sqrt{n}}, 50 + 1.96 \cdot (\frac{\sigma}{\sqrt{n}})]$

Generellt vill jeg ha 95% konfidens, vet konfidensintervallet

$\bar{x} \pm z_{\alpha/2}(\frac{\sigma}{\sqrt{n}})$ $\rightarrow$ Vanligt om jag behöver konfidens intervall

Exempel 6.1

Staten har en anslagsnorm om 30 elbilar finns en over utsläpping som kan antas vara normalfördelad med standardavvikelse 5 g/km. Ett stickprov av 100 elbilar har undersökts för gram. Intervallskattning under konfidensgrad 95 % ger det sanna medelvärdet

$$\bar{x} \pm 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

95% $\alpha = 5\%$

$26 \pm 1.96(\frac{5}{\sqrt{100}}) \rightarrow [25, 30.5]$

Sammanfattning

konfidens $\rightarrow$ sannolikheten att vi har rätt

sannolikhetsnivå $\rightarrow$ sannolikheten att vi har fel

S.sannolikhetsnivå betecknas med α

α = 1 - konfidensnivå

> Om $\bar{x}$ är normalfördelad till ett antal μ och en standardavvikelse σ kan man beräkna ett intervall som med sannolikhet $1-\alpha$ innehåller μ:
>
> $$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
>
> Det $1-\alpha$ kallas konfidensgrad och ofta brukar man ange det som en procentsiffra.
>
> Om $\bar{x}$ är normalfördelad till ett antal μ och en standardavvikelse σ kan man beräkna ett intervall som med sannolikhet $1-\alpha$ innehåller μ:
>
> $$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
>
> Det $1-\alpha$ kallas konfidensgrad och ofta brukar man ange det som en procentsiffra.

$\alpha = 5\%$

$1 - \alpha = 95\%$

$\bar{x} \pm z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$

konfidens = 99.5%

$\alpha = 0.5\%$

$\alpha = 0.002$

$\frac{\alpha}{2} = 0.25\% \rightarrow 0.0025$

$\bar{x} \pm z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.002}{2}} = Z_{0.001} \approx 3.08$

> Om $\bar{x}$ är normalfördelad till ett antal μ och en standardavvikelse σ kan man beräkna ett intervall som med sannolikhet $1-\alpha$ innehåller μ:
>
> $$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
>
> Det $1-\alpha$ kallas konfidensgrad och ofta brukar man ange det som en procentsiffra.

$B = z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$ $\leftarrow$ Vanligt om vi vill konfidensnivå behöver vidar ändra för bestämpunkt

Exempel 6.2

Staten har en anslagsnorm om 30 mikrofiber finns en over utsläpping som har antas vara normalfördelad med standardavvikelse 5 g/km. Ett stickprov av 100 elbilar har undersökts för gram och ett detta vilket blev vara approximativt normalfördelad. Intervallskattning under konfidensgrad 95 % ger det sanna medelvärdet av normalfördelad sanna medelvärdet för mikrofibrer. Vilk i hinda fallen med 99 % säkerhet.

$\bar{x} = 90$ $\sigma = 10$ $n = 16$

konfidens 99%

$\alpha = 1\%$

$\frac{\alpha}{2} = 0.005$

$z_{\frac{\alpha}{2}} = z_{0.005} = 2.58$

$B = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \rightarrow B = 2.58(\frac{10}{\sqrt{16}}) = 6.5$

99% $[90 - 6.5, 90 + 6.5] = [83.5, 96.5]$

Vanligt hinda kanske vi vilket man en stickprovsmedelvärdet av normalfördelad

$\rightarrow$ $\bar{x} \pm z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$

Men, hinda om vi inte hänger hinn man det säg om vi är säker på stickprov gå en approx normalfördelad population vi stickprov av stort (minst) min 5 stand $\rightarrow$ stickprovsmedelvärdet blir fördelad

> En $1-\alpha$ konfidensintervall för ytter μ är ordnad mer dismal av μ och stickprovsgrupp som ifrån en normalfördelad population.
>
> $$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$
>
> Det $1-\alpha$ konfidensintervallet ofta $\alpha = 1$ ta den beräknade ytter under lägre avsnitt i fördelningen beroende vilket $z_{\alpha/2}$ man i fördelningsfalet.

$\bar{x} \pm z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$

1. $n \geq 30$ $np \geq 5$, $n \geq 12$
2. σ okänt np
3. X approx normal np

Exempel 6.3

Staten har ingått ett slumpmässigt stickprov om 12 mikfiber finn sina kvinnliga borst. Stickprovet klassificeras av $\bar{x} = 90$ gram och $s = 32$ gram. Vilk i hinda fallen med 95 % säkerhet. Intervallskattning under konfidensgrad av vilken för kvällsbrun i hinna borst. Kvällsbrun vilk hun anses vara approximativt normalfördelad.

$\bar{x} = 90$, $s = 32$, $n = 12$

Canal. 95%: konfidens 95%, $\alpha = 5\%$, $\frac{\alpha}{2} = 0.025$, antal frihetsgrader $= 12 - 1 = 11$ ($t_{\frac{\alpha}{2}}$, df) $\rightarrow$ degrees freedom

$\bar{x} = t_{\frac{\alpha}{2}}(\frac{s}{\sqrt{n}})$

$90 \pm t_{0.025}(\frac{32}{\sqrt{12}}) = 90 \pm 2.201(\frac{32}{\sqrt{12}})$

95% $\rightarrow$ $[68.57, 110.58]$

Canal^2 99%: $\bar{x} = 90$, $s = 32$, $n = 12$, $df = 11$

konfidens 99%

$\alpha = 1\%$

$\frac{\alpha}{2} = 0.005$

$\bar{x} = t_{\frac{\alpha}{2}}(\frac{s}{\sqrt{n}})$

$90 \pm t_{0.005}(\frac{32}{\sqrt{12}}) = 90 \pm 3.106(\frac{32}{\sqrt{12}})$

99% $[61.5, 118.7]$

Hur konfidens vi (elmengenten) stödet?

$B = z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$

ex 95% konfidens, $\alpha = 5\%$, $\frac{\alpha}{2} = 0.025$

$\sigma = 10$, $n = 10$, $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$ $\Big\}$ $B = 1.96(\frac{10}{\sqrt{10}}) = 1.96$

ovan stickprovsmedelvärde $\bar{x}$ $[\bar{x} - 1.96, \bar{x} + 1.96]$ 95%

$B = 1.96$

$B = z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$ om vi ökar n, minskar B

$n = \left(z_{\frac{\alpha}{2}}\left(\frac{\sigma}{B}\right)\right)^2$

ex Vi vill ha 95% konfidens med en felmarginal $B = 0.5$. $\sigma = 10$. Fråga är: hur stort stickprov behöver jag ta för att uppnå det?

Dvs

95% $[\bar{x} - 0.5, \bar{x} + 0.5]$

$n = \left(1.96\left(\frac{10}{0.5}\right)\right)^2 = (1.96 \cdot 20)^2 = 1536.64 \approx \underline{1537}$

Det stickprovsstorlek du behöver (ti att uppnå $B = 0.5$ med 95%