Ex. 8.15

Stem-and-leaf diagram

2	9
3	8 5 9
4	2 1 7 8
5	1 3 5 6 7 5
6	1 8 1 0 2 5 7 3
7	9 7 3 2
8	1 0
9	2
10	0

x_8 x_{15} x_{16} x_{23}

(a) Antagandet om en normalfördelad population är
(- orimligt då bara heltal) efter skalering är värden "mindre diskreta"
- rimligt då formen passar

(b) Median: $\hat{x} = \dfrac{x_{15}+x_{16}}{2}$ där $(x_1, \dots, x_{30})$ är den ordnade datamängden

$= \dfrac{600+610}{2} = 605 = q_2$

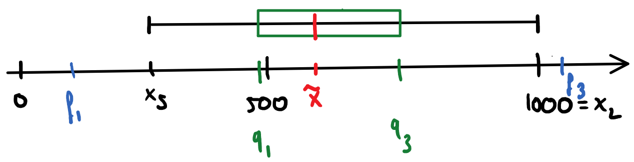
$q_1 = x_8 = 480$, $q_3 = x_{23} = 720$ (då $\frac{3}{4} \cdot 30 = 22.5$)

$iqr = q_3 - q_1 = 240$

$f_1 = q_1 - \frac{3}{2} \cdot (iqr) = 120$
$f_3 = q_3 + \frac{3}{2} \cdot (iqr) = 1080$ } "inner fences"

→ då $x_s = 290$ och $x_L = 1000$ ligger i $[120, 1080] = [f_1, f_3]$ finns det inga udda värden ("outliers").

range $= x_L - x_s = 710$



(c) 99% KI för μ, $\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} = z \sim N(0,1)$, $\dfrac{S^2(n-1)}{\sigma^2} \sim \chi^2_{(n-1)}$

okänd, men normalförd. population

$\dfrac{\bar{x}-\mu}{s/\sqrt{n}}$ är t-fördelad med $(n-1)$ frihetsgrader.

Här $\alpha = 0.01$, $n = 30$ och motsvarande K.I. ges av $\bar{x} \pm t_{\alpha/2} \dfrac{s}{\sqrt{n}}$

$t_{0.005} = 2.756$, $\bar{x} = \dfrac{1807}{3} \approx 602.3333$ och $s = 169.1089$

$$\left[602.3 - 2.756 \frac{169.1}{\sqrt{30}},\ 602.3 + 2.756 \frac{169.1}{\sqrt{30}}\right] = [517.2,\ 687.4]$$

(d) Minska konfidensnivån (alltså välja större α).
(eller ta större stickprov om möjligt)

8.32 $\mu = 0.6$

(a) $H_0 : \mu = \mu_0 = 0.6$ (eller ≥ 0.6) nollhypotesen

$H_1 : \mu < 0.6$ alternativet

(b) typ I-fel: förkastar H_0 trots att den stämmer

→ De nya motorer bedöms vara bättre trots att de inte är det.

typ II-fel: behåller H_0 trots att den inte stämmer

→ De nya motorer är bättre men det upptäcks inte av stickprovet.

(c) $n = 64$ stickprovets storlek, $\bar{x} = 0.5$ är stickprovsmedelvärdets utfall och $\sigma = 0.4$.

vad är P-värdet?

CGS: $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \approx Z \sim \mathcal{N}(0,1)$; $\quad z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{0.5-0.6}{0.4/\sqrt{64}} = -2$

$P = \mathbb{P}(Z \leq z_0) = 0.0228$

dvs. för $\alpha \geq P$ förkastas vi H_0, för $\alpha < P$ förkastar vi inte

Vi förkastar då P-värde är ganska litet.

→ Detta gör typ-I-fel möjligt (har dock bara slh. 0.0228).

Ex. 8.59 (a) $H_0: \mu = \mu_0 = 3$, $H_1: \mu < 3$.

(d) Det står inte i uppgiften, men vi antar att värden är normalfördelade, annars kommer vi inte vidare ($n=16$ är inte tillräckligt för CGS).

σ är okänd (om man inte vill använda skattningen $\hat{\sigma}$ från (c) som korrekt varians)

$\bar{x} = \frac{75}{32} \approx 2.344$. $s = 0.71458$ är utfallen på stickprovsmedelvärde och -varians.

$T := \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ är t-fördelad med $n-1=15$ frihetsgrader. $t_{1-\alpha} = t_{0.95} = -1.753$

- Utfallet $t_0 = \dfrac{\bar{x}-\mu_0}{s/\sqrt{n}} = -3.67 < t_{\alpha/2}$ → förkastar H_0.
- $H_0: \mu \leq \mu_0 = 2.5$, $H_1: \mu > 2.5$ leder till $t_\alpha = t_{0.05} = 1.753$ och $t_0 = -0.8746$

→ förkastar H_0 inte, markerads för [95% K.I. för μ: $\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = [1.963, 2.724]$ ligger till största del på rätt sida om 2.5].

Ex. 9.2 1 75 utav $n = 193$ icke längre fungerande protektorer var något mekaniskt fel anledningen.

(a) $X_j = \begin{cases} 1 & \text{om problemet i protektor } j \text{ har mekanisk orsak} \\ 0 & \text{annars} \end{cases}$ $\Rightarrow \hat{p} = \frac{1}{n}\sum_{j=1}^{n} X_j = \frac{75}{193} \approx 0.3886$.

(b) 95% K.I. ges av $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$; $\alpha/2 = 0.025$, $z_{\alpha/2} = 1.96$, $n = 193$

$0.389 \pm 1.96\sqrt{\dfrac{0.389 \cdot 0.621}{193}} = 0.389 \pm 0.069$.

(c) (se föreläsning 10) $n = \left(\dfrac{z_{\alpha/2}}{d}\right)^2 \cdot \hat{p}\cdot(1-\hat{p}) = 1014.14$ alltså ungefär 1015.
$d = 0.03$

Ex. 9.13

(a) $H_0: p \leq 0.6 = p_0$, $H_1: p > 0.6$

(b) Om nollhypotesen stämmer är

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$ approx. standard normalfördelad

Dvs $P(Z > z_0) = \alpha = 0.05$ för $z_0 = +1.645$

alltså $\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} > 1.645$ skulle leda till att man förkastar H_0.

(c) $\hat{p} = \frac{233}{375} \rightsquigarrow \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = 0.843$

-> vi förkastar H_0 inte ($z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} < 1.645$ ligger inte i "rejection region").

Vi kan alltså göra typ-II-fel, vilket betyder att vi behåller ett felaktigt antagande.

Ex. 9.19

$n_1 = 200$, $x_1 = 62$; $n_2 = 190$, $x_2 = 76$

(a) $\hat{p}_1 = \frac{x_1}{n_1} = 0.31$, $\hat{p}_2 = \frac{76}{190} = 0.4$ är (vvr. punktskattningar för p_1, p_2)

$\hat{p}_1 - \hat{p}_2 = -0.09$ är utfallet på (den vvr.) punktskattningen för $p_1 - p_2$

(b) 90% KI ges av

$$\underbrace{\hat{p}_1 - \hat{p}_2}_{=-0.09} \pm \underbrace{z_{\alpha/2}}_{= z_{0.05} = 1.645} \underbrace{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}_{\approx 0.0483}$$ alltså $[-0.169, -0.011]$

(c) Skulle vara förvånad då $p_1 - p_2 = 0$ inte ligger i 90% K.I. ovan.
(dvs. om det stämmer skulle slh. för att få en så stor differens mellan $\hat{p}_1$ och $\hat{p}_2$ (eller större) vara < 0.1).