Hypotestest för μ när σ är okänd, t-test

Vad vi har gått igenom hittills om hypotestest fungerar om n är stort nog eller om stickprovet är normalfördelat. Vad händer om vi inte känner till σ? Om n är stort nog så ersätter vi σ med dess punktskattare s. Teststatistikan blir då $Z_i = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ som enligt centrala gränsvärdessatsen är normalfördelad. Vi fortsätter precis som tidigare!

Om σ är okänd och om n inte är stort nog så måste vi anta att stickprovet är normalfördelat. Teststatistikan blir igen $Z_i = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ men nu är den t-fördelad med $n-1$ frihetsgrader. T-test

Vi gör en tabell för att sammanfatta de olika fallen vi har gått igenom hittills. ~~Tabellen gäller om vi vill testa en nollhypotes på formen $H_0: \mu = \mu_0$~~

Populationen normalfördelad?	σ känd?	teststatistika	fördelning
Ja	Ja	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$N(0,1)$
Ja	Nej	$T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$	t_{n-1}
Nej	Ja	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	(om n stort) $N(0,1)$
Nej	Nej	$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}}$	(om n stort) $N(0,1)$

Om populationen inte kan antas vara normalfördelad och om n inte är stort nog så undviker vi att utföra hypotestest!

Hypotestest för en proportion. (kap. 8.6)

Precis som för konfidensintervall så kan man modifiera hypotestest för ett väntevärde något så att analysen kan göras för en proportion istället. Hypoteserna kan t. ex. se ut så här (för tvåsidigt).

$$H_0: p = p_0$$
$$H_1: p \neq p_0$$

Teststatistikan blir här ($\hat{p}$ = punktskattaren)

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

som är $N(0,1)$-fördelad om n är stort.

Tumregel: $np_0 \geq 15$ och $n(1-p_0) \geq 15$

Övning 87: En undersökning visade att 401 av 835 tillfrågade (amerikanska) ungdomar växte upp i ett hushåll med endast en förälder. Baserat på denna information, kan du dra slutsatsen att mer än 45% av amerikanska ungdomar har växt upp i ett sådant hushåll? Testa på signifikansnivå $\alpha = 0.05$.

Lösning: Låt p vara den riktiga proportionen av amerikanska ungdomar som växt upp i enförälderhem.

Hypoteserna blir nu

$$H_0 : p \leq 0.45$$

$$H_1 : p > 0.45 .$$

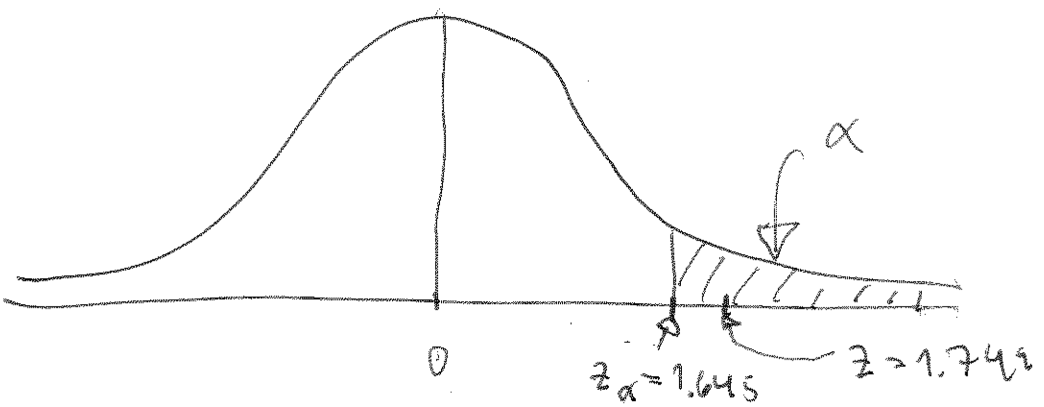
Kom ihåg att det är $p > 0.45$ vi vill statistiskt säkerställa, alltså blir det alternativhypotesen. Teststatistikan blir

$$\cancel{Z} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \quad \text{vilken är } N(0,1)\text{-fördelad.}$$

Eftersom $\hat{p} = \frac{401}{835} = 0.4802$ blir den observerade teststatistikan

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.4802 - 0.45}{\sqrt{\frac{0.45(1-0.45)}{835}}} = 1.749$$

Vi vill alltså jämföra detta värde med det kritiska värdet z_α. Tabellerna ger att $z_{0.05} = 1.645$. Eftersom $1.749 > 1.645$ så förkastar vi H_0 på signifikansnivå $\alpha = 0.05$



Slutsats: Ja, vi kan dra slutsatsen att mer än ③
45% av amerikanska ungdomar växt upp i ~~ett~~
~~ett~~ hushåll med en förälder.