Hypotesetesting.

Dekkes av pensumsidene i kap. 10 og forelesingsnotatene.

Hypotesetesting er en systematisk fremgangsmåte for å undersøke hypoteser (påstander) knyttet til parametre i sannsynlighetsfordelinger.

Eksempel: Bærebjelker. Bæreevnen antas normalfordelt med forventning μ og varians σ^2. Før vi eventuelt tar i bruk bjelker av denne typen stiller vi som krav at $\mu > 5.5$ og $\sigma^2 < 0.25$. Tyder målingene 5.9, 6.1, 5.4, 5.9, 5.7, 6.5, 5.8, 6.0, 5.3, 6.3, 5.8, 5.8, 5.8 og 6.0 på at dette er oppfylt? $\square$

Eksempel: Levetid til ny type kretskort. Levetiden antas eksponensialfordelt med parameter β. Et mål i utviklingen av kretskortet var at det skulle ha en forventet levetid på mer enn 700. Tyder de observerte tidene 612, 1009, 95, 1303, 599, og 780 på at dette er oppfylt? $\square$

Formulere en hypotesetest:

Når vi skal test om en parameter θ kan påstås å være større enn en verdi θ_0 skriver vi dette som:

$$H_0: \quad \theta = \theta_0 \qquad \text{mot} \qquad H_1: \quad \theta > \theta_0$$

Når vi skal test om $\theta < \theta_0$ skriver vi:

$$H_0: \quad \theta = \theta_0 \qquad \text{mot} \qquad H_1: \quad \theta < \theta_0$$

Når vi skal test om $\theta \neq \theta_0$ skriver vi:

$$H_0: \quad \theta = \theta_0 \qquad \text{mot} \qquad H_1: \quad \theta \neq \theta_0$$

Merk at H_1 er påstanden vi skal undersøke om er riktig.

Prinsipp: Konservativ fremgangsmåte - ønsker ikke å påstå at det vi skal undersøke om er tilfelle, H_1, virkelig er tilfelle med mindre dataene peker klart i denne retningen.

Vi antar derfor i utgangspunktet at H_0 er korrekt, og gjør all regning under denne antagelsen. Påstår bare H_1 dersom de observerte dataene er lite sannsynlige under H_0-modellen og peker klart i retning av H_1-modellen.

Setter derfor $\alpha = P(\text{forkaste } H_0 | H_0 \text{ er rett})$ liten.

α kalles *(signifikans-)nivået* til testen.

NB! Resultatet av en hypotesetest er:

- Enten at vi forkaster H_0 - som vil si at vi påstår at H_1 er rett.
- Eller at vi ikke forkaster H_0 - som vil si at <u>både H_0 og H_1 er mulige</u> (H_1 er fremdeles mulig selv om dataene ikke peker sterkt i den retning!).

Hypotesetesting

(I de situasjonene vi skal se på i dette kurset)

$X_1, \ldots, X_n$ u.i.f. $f(x;\theta)$.

$$H_0: \quad \theta = \theta_0 \qquad \text{mot} \qquad H_1: \quad \begin{array}{ll} \text{a)} & \theta > \theta_0 \\ \text{b)} & \theta < \theta_0 \\ \text{c)} & \theta \neq \theta_0 \end{array}$$

1. Finn estimator for θ: $\hat{\Theta}$.
2. Finn en testobservator $W = h(\hat{\Theta}, \theta_0)$ som har en kjent fordeling under H_0.
3. Dersom testobservatoren $W = h(\hat{\Theta}, \theta_0)$ er stor når $\hat{\Theta}$ er stor og liten når $\hat{\Theta}$ er liten så:

   a) Forkaster vi H_0 dersom $W \geq w_\alpha$
   (der $P(\text{forkaste } H_0 | H_0) = P(W \geq w_\alpha | H_0) = \alpha$).

   b) Forkaster vi H_0 dersom $W \leq w_{1-\alpha}$
   (der $P(\text{forkaste } H_0 | H_0) = P(W \leq w_{1-\alpha} | H_0) = \alpha$).

   c) Forkaster vi H_0 dersom $W \leq w_{1-\alpha/2}$ *eller* $W \geq w_{\alpha/2}$ (der $P(\text{forkaste } H_0 | H_0) = P(W \geq w_{\alpha/2} | H_0) + P(W \leq w_{1-\alpha/2} | H_0) = \alpha$).

Eksempel: Bæreevne til bærebjelker. $X_1, \ldots, X_n$ u.i.f. $N(\mu, \sigma^2)$. Både μ og σ^2 er ukjente. Vi ønsker å teste om $\mu > 5.5$ og om $\sigma^2 < 0.25$.

Test for μ:

$$H_0: \quad \mu = 5.5 \qquad \text{mot} \qquad H_1: \quad \mu > 5.5$$

$$\text{Estimator:} \qquad \hat{\mu} = \bar{X}$$

$$\text{Testobservator:} \qquad T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1} \quad \text{under } H_0$$

Vi forkaster H_0 dersom $T \geq t_{\alpha,n-1}$.

Velger vi $\alpha = 0.05$ får vi med $n = 14$ at vi forkaster H_0 dersom $T \geq t_{0.05,13} = 1.771$.

Med de observerte dataene får vi $\bar{x} = 5.8786$ og $s = \sqrt{0.0972} = 0.312$ som gir

$$t = \frac{5.8786 - 5.5}{0.312/\sqrt{14}} = 4.54 > t_{0.05,13} = 1.771$$

dvs vi forkaster H_0 på 5% nivå. Dataene gir grunnlag for å hevde at $\mu > 5.5$.

Dvs $\hat{\mu} = \bar{x} = 5.8786$ er så mye større enn 5.5 at det gir grunnlag for å hevde at $\mu > 5.5$

Test for σ^2:

$$H_0: \ \sigma^2 = 0.25 \qquad \text{mot} \qquad H_1: \ \sigma^2 < 0.25$$

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

Teorem 8.4/tabell s. 27 gir at:

$$(n-1)\frac{S^2}{\sigma_0^2} = \sum_{i=1}^{n}\frac{(X_i - \bar{X})^2}{\sigma_0^2} \sim \chi^2_{n-1} \quad \text{under } H_0$$

Vi forkaster H_0 dersom $(n-1)\frac{S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha,n-1}$.

Med $\alpha = 0.05$ og $n = 14$ blir $\chi^2_{1-\alpha,n-1} = \chi^2_{0.95,13} = 5.892$

Observert:

$$(n-1)\frac{s^2}{\sigma_0^2} = (14-1)\frac{0.0972}{0.25} = 5.05 < \chi^2_{0.95,13} = 5.892$$

dvs vi forkaster H_0 på 5% nivå. Dataene gir grunnlag for å hevde at $\sigma^2 < 0.25$.

Dvs $\hat{\sigma}^2 = S^2 = 0.0972$ er så mye mindre enn 0.25 at det gir grunnlag for å hevde at $\sigma^2 < 0.25$.

$\square$

Eksempel: Levetid kretskort. Levetiden antas eksponensialfordelt med forventning β. Ønsker å teste om målet $\beta > 700$ ser ut til å være oppfylt.

$$H_0: \ \beta = 700 \qquad \text{mot} \qquad H_1: \ \beta > 700$$

Estimator:
$$\hat{\beta} = \bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$$

Fra tidligere (se notatene om estimering) får vi:

$$\frac{2n}{\beta_0}\hat{\beta} \sim \chi^2_{2n} \quad \text{under } H_0$$

Vi forkaster H_0 dersom $\frac{2n}{\beta_0}\hat{\beta} \geq \chi^2_{\alpha,2n}$.

Med $\alpha = 0.05$ og $n = 6$ blir $\chi^2_{\alpha,2n} = \chi^2_{0.05,12} = 21.026$

Observert:

$$\frac{2n}{\beta_0}\hat{\beta} = \frac{12}{700} \cdot 733 = 12.57 < \chi^2_{0.05,12} = 21.026$$

dvs vi <u>forkaster ikke H_0</u> på 5% nivå. Dataene gir ikke grunnlag for å hevde at $\beta > 700$. □

Eksempel: I en produksjonsprosess har det tidligere blitt produsert 18% B-varer. Etter en omlegging av rutinene lurer man på om andelen B-varer har *endret* seg. Blant 75 varer produsert etter omleggingen blir 17 klassifisert som B-varer. Tyder dette på at andelen B-varer har endret seg?

Merk først: X="antall B-varer blant n tilfeldig valgte varer" vil være $B(n, p)$-fordelt, der p er sannsynligheten for at en tilfeldig vare er B-vare (=andel B-varer).

$$H_0: \quad p = 0.18 \qquad \text{mot} \qquad H_1: \quad p \neq 0.18$$

$$\text{Estimator:} \qquad \hat{p} = \frac{X}{n}$$

Fordeling til $\hat{p}$?

Fra kap. 6 har vi at når $X \sim B(n, p)$ og $np > 5$ og $n(1-p) > 5$ så er $X \approx N(np, np(1-p))$. Dette er oppfylt her, og når X er tilnærmet normalfordelt vil også $\hat{p} = \frac{X}{n}$ være tilnærmet normalfordelt.

$$
\begin{aligned}
\mathrm{E}(\hat{p}) &= \mathrm{E}(\frac{X}{n}) = \frac{1}{n}\mathrm{E}(X) = \frac{1}{n}np = p \\
\mathrm{Var}(\hat{p}) &= \frac{1}{n^2}\mathrm{Var}(X) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n} \\
\Rightarrow \quad Z &= \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \approx N(0,1) \quad \text{under } H_0
\end{aligned}
$$

Vi forkaster H_0 dersom $Z \leq -z_{\alpha/2}$ eller $Z \geq z_{\alpha/2}$.

Velger vi $\alpha = 0.05$ får vi $z_{\alpha/2} = z_{0.025} = 1.96$.

Observerte data gir $\hat{p} = 17/75 = 0.227$ og dermed:

$$
z = \frac{0.227 - 0.18}{\sqrt{0.18(1-0.18)/75}} = 1.06
$$

dvs vi forkaster ikke H_0 på 5% nivå. Dataene gir ikke grunnlag for å hevde at $p \neq 0.18$. $\square$

Sammenheng mellom (tosidig) hypotesetesting og konfidensintervall: Les selv i boka, kap. 10.6!