SF1901 Sannolikhetsteori och statistik I

Jimmy Olsson

Föreläsning 14
7 december 2017

Idag

Förra gången

Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Idag

Förra gången

Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Förra gången: hypotesprövning

- Vi har ett slumpmässigt stickprov $\mathbf{x} = (x_1, x_2, \ldots, \; x_n)$ som anses vara ett utfall av s.v. $\mathbf{X} = (X_1, X_2, \ldots, \; X_n)$ med någon fördelning.
- Vi vill testa en viss *nollhypotes* H_0 rörande fördelningen.
- För att pröva H_0 definierar vi först en *testvariabel* $t_{\text{obs}} = t(\mathbf{x})$ som är en observation av motsvarande *stickprovsvariabel* $t(\mathbf{X})$
- och definierar en *signifikansnivå* (eller *felrisk*) $0 < \alpha < 1$.

Generell metod (forts.)

- Testet utförs nu som följer:

$$\text{om} \begin{cases} t_{\text{obs}} \in \mathrm{C}, & \text{förkasta } H_0, \\ t_{\text{obs}} \notin \mathrm{C}, & \text{förkasta ej } H_0, \end{cases}$$

  där det *kritiska området* C är avpassat så att

$$\mathbb{P}(H_0 \text{ förkastas}) = \mathbb{P}(t(\mathbf{X}) \in \mathrm{C}) = \alpha \quad \text{om } H_0 \text{ är sann.}$$

- Att $t_{\text{obs}} \in \mathrm{C}$ är alltså en extrem händelse om H_0 är sann, och vi kan alltså förkasta H_0 med gott samvete om detta inträffar.
- Om H_0 förkastas sägs det *föreligga en signifikant avvikelse från nollhypotesen H_0 på nivån α*.

Generell metod (forts.)

- Vi kommer i regel att låta H_0 vara en *enkel hypotes*, alltså omfatta endast *ett* parametervärde θ_0, dvs.

$$H_0 : \theta = \theta_0.$$

- För att rationellt kunna välja kritiskt område C specificeras en *mothypotes* H_1 som kan (men inte måste) vara sann om H_0 inte är det.
- Mothypotesen H_1 kan vara *sammansatt*, dvs. omfatta fler än ett parametervärde.
- Ett bra test förkastar H_0 med stor sannolikhet om H_1 är sann. Denna egenskap beskrivs av testets *styrka*, vilken kommer att diskuteras senare idag.

Idag

Förra gången

Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Metoder för hypotesprövning

- Den beskrivna metoden implementeras ofta i form av
  - *direktmetoden*,
  - *konfidensmetoden*.

Direktmetoden

- Låt $\mathbb{P}_\theta$ beteckna sannolikheter då θ är rätt parameter.
- Med *direktmetoden* beräknas *P-värdet*

$$P = \mathbb{P}_{\theta_0}(t(\mathbf{X}) \geq t_{\text{obs}}) \text{ (eller ev. } = \mathbb{P}_{\theta_0}(t(\mathbf{X}) \leq t_{\text{obs}}))$$

  baserat på den observerade testvariabeln t_{obs}.
- P-värdet tolkas som ''sannolikheten, under H_0, att erhålla den observerade testvariabeln eller något än mer extremt utfall''.
- Om $P \leq \alpha$ förkastas H_0,
- vilket i termer av allmänna metoden betyder att t_{obs} ligger i det kritiska området C.

Exempel: skev tärning?

- En (ev. skev) tärning ger en 6:a med sannolikheten p. Man vill pröva hypotesen

$$H_0 : p = p_0 = \frac{1}{6} \quad \text{mot} \quad H_1 : p \neq \frac{1}{6}$$

  på signifikansnivån $\alpha = 5\%$.
- Tärningen kastas $n = 18$ ggr, varav 6 st. 6:or erhålls. Kan H_0 förkastas till förmån för H_1 på den valda nivån?

[svar: Nej! P-värde: 0.1028.]

Konfidensmetoden

- Med *konfidensmetoden* konstrueras ett $(1-\alpha)$-gradigt konfidensintervall
$$I_\theta = (a_1(\mathbf{x}), a_2(\mathbf{x}))$$
för parametern θ.
- Om $\theta_0 \notin I_\theta$ förkastas H_0.
- Formen på intervallet (en- eller tvåsidigt, uppåt eller nedåt begränsat) bestäms av mothypotesen H_1:
$$H_1 : \theta > \theta_0 \Rightarrow \text{ensidigt, nedåt bergänsat,}$$
$$H_1 : \theta < \theta_0 \Rightarrow \text{ensidigt, uppåt bergänsat,}$$
$$H_1 : \theta \neq \theta_0 \Rightarrow \text{tvåsidigt.}$$
- Se exempel från tidigare föreläsningar!

Konfidensmetoden (forts.)

- I termer av allmänna metoden betyder detta att

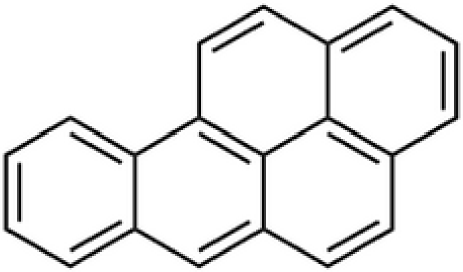
$$C = \{\mathbf{x} : \theta_0 \le a_1(\mathbf{x}) \text{ eller } \theta_0 \ge a_2(\mathbf{x})\},$$

  så att

$$\mathbb{P}_{\theta_0}(H_0 \text{ förkastas}) = \mathbb{P}_{\theta_0}(\mathbf{X} \in C) = \alpha.$$

- Med andra ord: om intervallet inte täcker θ_0 kan vi förkasta H_0 med gott samvete.

Exempel: förorenad jord? — igen



Figur: Bensopyren

Exempel: förorenad jord? — igen (forts.)

- Man planerar att bygga bostäder på mark där det tidigare bedrivits industriverksamhet och vill därför undersöka förekomsten av PAH* i jordlagret. Naturvårdsverkets PAH-riktvärde för känslig markanvändning är 0.3 mg/kg jord. Med hjälp av skruvborrning erhålls ett stickprov omfattande n jordprover med representativ halt $\bar{x}$. Mätmetodens standardavvikelse anses vara känd, $\sigma = 0.05$ mg/kg.
  (a) Utforma ett lämpligt test på nivån 5% för om PAH-halten μ i jordlagret överstiger Naturvårdsverkets riktvärde. Antag normalfördelning. Ange klart noll- och mothypotes samt kriterium för när nollhypotesen skall förkastas.

$$\left[\text{svar: } H_0 \text{ skall förkastas om } \bar{x} > 0.3 + \lambda_{0.05}\frac{0.05}{\sqrt{n}}.\right]$$

*Polycykliska aromatiska kolväten. Exponering för dessa kan medföra ökad risk för cancer.

Idag

Förra gången

Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Fel av typ I och II

- När vi utför en hypotesprövning riskerar vi att göra två typer av felslut, nämligen
  (I) förkasta H_0 trots att H_0 är sann (*typ I-felet* eller α-*felet*),
  (II) inte förkasta H_0 till förmån för H_1 trots att H_1 är sann (*typ II-felet* eller β-*felet*).
- Risken för typ I-felet ges av testets felrisk α, och genom att minska det kritiska området (dvs. höja tröskeln för att förkasta H_0) kan denna risk minskas.
- Minskad typ I-risk ger dock ökad typ II-risk.
- Ett bra test har låga typ I- och typ II-risker.

Styrkefunktionen

- Kom ihåg att $\mathbb{P}_\theta$ betecknar sannolikheter då θ är rätt parameter.

Definition

Styrkefunktionen för ett hypotestest ges av

$$h(\theta) = \mathbb{P}_\theta(H_0 \text{ förkastas}).$$

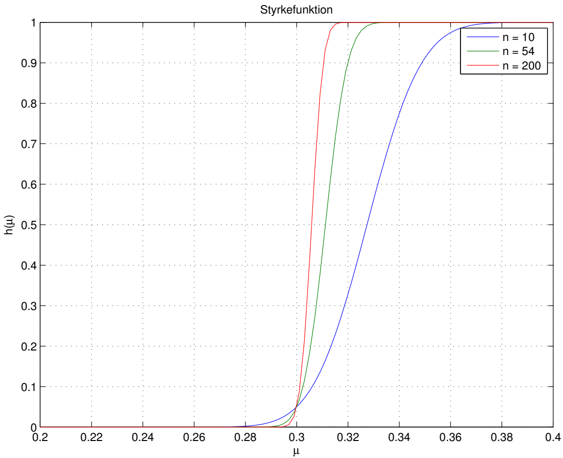
- Ett test är bra om $h(\theta)$ är
  - *liten* (dvs. nära 0) för alla $\theta \in H_0$ (låg typ I-risk),
  - *stor* (dvs. nära 1) för alla $\theta \in H_1$ (låg typ II-risk).

Exempel: förorenad jord? — igen (forts.)

- Man planerar att bygga bostäder på mark där det tidigare bedrivits industriverksamhet och vill därför undersöka förekomsten av PAH i jordlagret. Naturvårdsverkets PAH-riktvärde för känslig markanvändning är 0.3 mg/kg jord. Med hjälp av skruvborrning erhålls ett stickprov omfattande n jordprover med representativ halt $\bar{x}$. Mätmetodens standardavvikelse anses vara känd, $\sigma = 0.05$ mg/kg.
  (b) Hur många jordprover n bör man ta för att testets styrka i $\mu = 0.32$ mg/kg skall vara 90%?

[svar: $n = 54$ st.]

Exempel: förorenad jord? — igen (forts.)



Figur: Styrkefunktioner $h(\mu)$ för olika värden på n. Notera att $h(0.3) = 0.05 = \alpha$ for alla n.

Idag

Förra gången

Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Användning av normalapproximation

- Om H_0 rör en parameter θ för vilken man har en approximativt normalfördelad skattning $\theta^* \in \mathrm{AsN}(\theta, D)$ ger det föregående resonemanget ett test med *approximativ* signifikansnivå.
- Vi förkastar alltså $H_0 : \theta = \theta_0$ om den under H_0 approximativt normalfördelade testvariabeln

$$\tau = \frac{\theta^* - \theta_0}{D}$$

  överskrider något extremt värde som ges av en normalfördelningskvantil.
- Ofta beror standardavvikelsen D på θ, och under H_0 skall således $D = D(\theta_0)$ användas.

Exempel: skev tärning? — igen

- En (ev. skev) tärning ger en 6:a med sannolikheten p. Man vill pröva

$$H_0 : p = p_0 = \frac{1}{6} \quad \text{mot} \quad H_1 : p \neq \frac{1}{6}$$

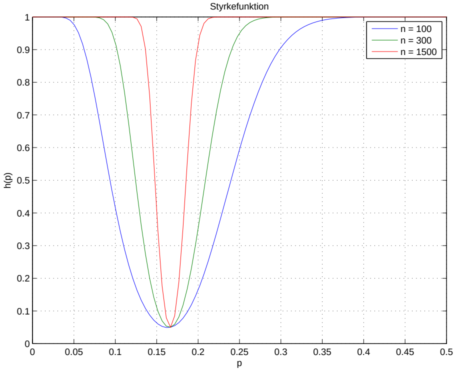
på signifikansnivån $\alpha = 0.05$.

(a) Tärningen kastas $n = 100$ ggr, varav 26 st. 6:or erhålls. Kan H_0 förkastas till förmån för H_1 på den valda nivån? Inför lämpliga approximationer om nödvändigt.

[svar: Ja!]

(b) Bestäm testets styrkefunktion $h(p)$.

Exempel: skev tärning? — igen (forts.)



Figur: Styrkefunktioner $h(p)$ för olika värden på n.

Idag

Förra gången

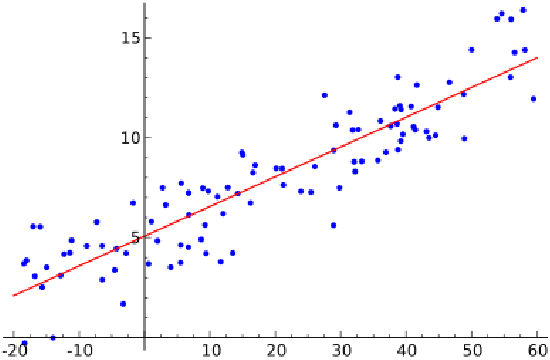
Direkt- och konfidensmetoderna

Styrkefunktionen (Kap 13.4)

Användning av normalapproximation (Kap. 13.7)

Regressionsanalys (Kap. 14)

Modell



Figur: I en linjär regressionsmodell anpassas bästa möjliga räta linje efter data.

Modell (forts.)

- Det föreligger n par av värden

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

där x_1, x_2, ..., x_n är fixa tal (*regressionsvariabler*) och y_1, y_2, ..., y_n är observationer av

$$Y_i = \underbrace{\alpha + \beta x_i}_{\overset{\text{not.}}{=} \mu_i} + \varepsilon_i, \quad i = 1, 2, \ldots, n,$$

med oberoende *residualer* $\varepsilon_i \in \mathrm{N}(0, \sigma)$.

- Linjen $y = \alpha + \beta x$ kallas den *teoretiska regressionslinjen* och visar hur väntevärdet för Y_i:na beror av x.

ML-skattningar av α, β och σ (forts.)

- Parametrarna α, β och σ är okända. Det visar sig att de korrigerade ML-skattningarna ges av

$$\beta^*_{\text{obs}} = \frac{s_{xy}}{s_{xx}}, \quad \alpha^*_{\text{obs}} = \bar{y} - \beta^*_{\text{obs}}\bar{x}, \quad (\sigma^2)^*_{\text{obs}} = \frac{Q_0}{n-2},$$

där

$$s_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), \quad s_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2,$$

$$Q_0 = Q(\alpha^*_{\text{obs}}, \beta^*_{\text{obs}}) = \sum_{i=1}^{n}(y_i - \alpha^*_{\text{obs}} - \beta^*_{\text{obs}}x_i)^2 = s_{yy} - \frac{s_{xy}^2}{s_{xx}}.$$

- Skattningarna av α och β är desamma även om σ är känd.

Fördelningar hos skattningar

- Det är viktigt att notera att α^* och β^* är *linjära* i y_i:na.
- Detta betyder att dessa skattningar är *normalfördelade*.
- Man visar t.ex. (övning!) att

$$\alpha^* \in \mathrm{N}\left(\alpha, \sigma\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{s_{xx}}}\right), \quad \beta^* \in \mathrm{N}\left(\beta, \frac{\sigma}{\sqrt{s_{xx}}}\right).$$

- Dessutom kan man visa att

$$\frac{(n-2)S^2}{\sigma^2} \in \chi^2(n-2).$$

- Slutligen gäller att S^2 är *oberoende* av α^* och β^*.

Konfidensintervall (Kap. 14.4)

- Genom att tillämpa λ- eller t-metoden erhålls direkt t.ex. det tvåsidiga, $(1-\alpha)$-gradiga konfidensintervallet

$$I_\beta = \begin{cases} (\beta^*_{\text{obs}} \pm \lambda_{\alpha/2} D) & \text{om } \sigma \text{ är känt } (D = \sigma/\sqrt{s_{xx}}), \\ (\beta^*_{\text{obs}} \pm t_{\alpha/2}(n-2) d) & \text{om } \sigma \text{ är okänt } (d = s/\sqrt{s_{xx}}) \end{cases}$$

  för lutningen β.
- Detta kan användas för att med hjälp av konfidensmetoden pröva hypotesen att $\beta = 0$, dvs. undersöka om y överhuvudtaget beror på x.

Skattning av punkt på linjen

- Givet något värde $x = x_0$ är vi ofta intresserade av att skatta väntevärdet
$$\mu_0 = \alpha + \beta x_0,$$
dvs. motsvarande punkt på den teoretiska regressionslinjen.
- Som skattning tar man helt enkelt
$$\mu_0^* = \alpha^* + \beta^* x_0 \overset{\text{övning!}}{\in} \mathrm{N}\left(\mu_0, \sigma\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{s_{xx}}}\right).$$
- Konfidensintervall erhålls i sedvanlig ordning med λ- eller t-metoden (beroende på om σ är känt eller inte).
- Se FS!

Utvidgningar

- *Multipel regression* (se Kap. 14.5):

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon,$$

- *generella linjära modeller*,
- *generaliserade linjära modeller* (t.ex. *logistisk regression*),
- ...
- SF2930 Regressionsanalys, 7.5 hp!

Nästa föreläsning

- Test av fördelning,
- homogenitetstest,
- oberoendetest.