Matematisk statistik för B, K, N, BME och Kemister

Föreläsning 6

Johan Lindström

13 september 2017

Stickprov (Kap. 7.1)

Ett **stickprov**, $x_1, x_2, \ldots, x_n$, är **observationer** av s.v. $X_1, \ldots, X_n$ från någon fördelning $X_i \in F(\vartheta)$ där ϑ är en okänd **parameter**.

Skattning (Kap. 7.1)

En **skattning** av ϑ, $\vartheta^*(x_1, \ldots, x_n)$ är en observation av den s.v. $\vartheta^*(X_1, \ldots, X_n)$. Båda betecknas oftast bara med ϑ^*.

En skattning är:

1. En funktion från observationer till ett tal.
2. En stokastisk variabel.
3. Ett tal.

Konfidensintervall för μ i $\mathsf{N}\left(\mu, \sigma^2\right)$ (Kap. 7.3)

$x_1, \ldots, x_n$ observationer av $X_i \in \mathsf{N}\left(\mu, \sigma^2\right)$
σ^2 **känd:**

$$I_\mu = \bar{x} \pm \lambda_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \mu^* \pm \lambda_{\alpha/2} D(\mu^*)$$

σ^2 **okänd:**

$$I_\mu = \bar{x} \pm t_{\alpha/2}(n-1) \frac{s}{\sqrt{n}} = \mu^* \pm t_{\alpha/2}(f) d(\mu^*)$$

Där kvantilerna ges av:

- $\lambda_{\alpha/2}$ är $N(0,1)$-fördelningens $\alpha/2$-kvantil (Tabell 2)
- $t_{\alpha/2}(n-1)$ är t-fördelningens $\alpha/2$-kvantil (Tabell 3)

Konfidensintervall för σ^2 i $\mathsf{N}\left(\mu, \sigma^2\right)$ (Kap. 8.1.2)

$x_1, \ldots, x_n$ observationer av $X_i \in \mathsf{N}\left(\mu, \sigma^2\right)$
Ett $1-\alpha$ konfidensintervall för σ^2 ges av

$$I_\sigma^2 = \left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} \right)$$

Där

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

och $\chi^2_{\alpha/2}(n-1)$ är χ^2-fördelningens kvantiler (Tabell 4)

Ex: Järnhalt i jord (Två Stickprov) (Kap. 7.7)

En geokemist är intresserad av att undersöka hur halten av järn varierar i skogsmark. På två lokaler gräver hon 6 resp. 7 gropar, tar ett prov från varje grop och bestämmer sedan järnhalten (mg/g):

Lokal I, x_i:	23.3	9.0	9.8	19.9	15.0	20.5	
Lokal II, y_i:	20.8	26.5	18.3	28.6	33.1	21.5	29.3

Ange en lämplig modell och gör ett konfidensintervall för skillnaden i järnhalt mellan de två lokalerna?

Sammanvägd variansskattning (Kap. 7.4)

Om vi har

$$x_1, \ldots, x_{n_x} \qquad \text{obs. av } X_i \in \mathrm{N}\left(\mu_x, \sigma^2\right)$$

$$y_1, \ldots, y_{n_y} \qquad \text{obs. av } Y_i \in \mathrm{N}\left(\mu_y, \sigma^2\right)$$

kan den gemensamma variansen σ^2 skattas med

$$s_p^2 = \frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x - 1 + n_y - 1} = \frac{Q}{f}, \qquad \left(\frac{Q}{\sigma^2} \in \chi^2(f)\right)$$

Stickprov i par (Kap. 7.8)

Vid många mätsituationer är det vanligt att man mäter före och efter en "behandling" på n inbördes olika föremål.

Modell:
Före: $X_i \in \mathsf{N}\left(\mu_i, \sigma_1^2\right)$
Efter: $Y_i \in \mathsf{N}\left(\mu_i + \Delta, \sigma_2^2\right)$

Vi vill nu skatta effekten av behandlingen (Δ). Bilda $Z_i = Y_i - X_i \in \mathsf{N}\left(\Delta, \sigma^2\right)$.

Stickprov i par eller två stickprov?

- Blodtrycket hos ett antal patienter mäts förre och efter behandling med blodtryckssänkande medicin; konfidensintervall för sänkningen?
- Luftkvaliteten mäts längs Hornsgatan i Stockholm vintern 2009 (dubbdäck fortfarande tillåtna) och 2010 (efter dubbdäcksförbud); konfidensintervall för skillnaden i luftkvalitet?
- pH-värdet möts varje dag i Höjeå förre och efter Lunds reningsverk; konfidensintervall för skillnaden?

Ex: Järnhalt i jord (Stickprov i par)

Geokemisten är speciellt intresserad av eventuella skillnader i järnhalten mellan olika nivåer i groparna. Hon tar därför, från 4 olika gropar, ett prov på A-nivå (nära ytan) och ett prov på C-nivå (c:a 1 meter djupt). Området är av mycket heterogen karaktär, dvs troligen varierar järnhalten mycket mellan olika gropar.

Grop nr:	1	2	3	4
Nivå A:	19.15	23.35	20.1	16.7
Nivå C:	21.96	27.7	22.93	19.02
$z_i = y_i - x_i$	2.81	4.35	2.83	2.32

Ange en lämplig modell och gör ett konfidensintervall för skillnaden i järnhalt mellan A- och C-nivårena.

Repetition
- Skattning
- Konfidensintervall

Konfidensintervall
- Två Stickprov
- Stickprov i par

Hypotesprövning
- Metoder

Styrkefunktion
- Ex: Rattonykterhet

Statistikteori – översikt

Punktskattning
Hur gör man en bra gissning av en okänd storhet? Hur vet man att den är bra?

Intervallskattning
Hitta istället ett intervall som täcker den okända storheten med en given (stor) sannolikhet.

Hypotestest
Om gissningen blev 0.013, kan rätt värde på den okända storheten ändå vara 0.01?

Regression
Hur vet vi om två variabler påverkar varandra?

Försöksplanering & Faktorförsök
Hur konstruerar man studier som på bäst sätt (minst antal mätningar) undersöker effekten av olika faktorer (behandlingar)?

Hypotesprövning (Kap. 7.5)

H_0 **förkastas** om observationerna, ϑ^*, avviker för mycket från **nollhypotesen** ϑ_0.

Testa **nollhypotesen** $H_0: \ \vartheta = \vartheta_0$
mot **mothypotesen** (tex) $H_1: \ \vartheta \neq \vartheta_0$

på nivån α; **felrisken** α ges av

$$\alpha = \mathrm{P}(H_0 \text{ förkastas trots att den är sann})$$

"För mycket" beror på osäkerheten i skattningen, $\mathrm{D}(\vartheta^*)$, samt felrisken α.
Normalt är $\alpha = 0.05$, 0.01 eller 0.001.

Olika metoder för att utföra hypotestest

1. **Direktmetoden eller P-värde** **(Def. 7.35)**
   - Antag att H_0 är sann
   - Räkna ut **P-värdet** $p = \mathrm{P}(\text{Få det vi fått eller värre})$
   - Om $p < \alpha$ förkastas H_0
2. **Konfidensmetoden** **(Kap. 7.5.2)**
   Gör ett konfidensintervall med konfidensgraden $1-\alpha$ och förkasta H_0 på nivån α om intervallet ej täcker ϑ_0.
   Intervallen skall, beroende på H_1, vara

Test	$H_1: \vartheta < \vartheta_0$	$H_1: \vartheta \neq \vartheta_0$	$H_1: \vartheta > \vartheta_0$
Intervall:	uppåt begr	tvåsidigt	nedåt begr

3. **Testkvantitet $T(X)$ och kritiskt område C** **(Kap. 7.5.3)**
   Förkasta H_0 om testskvantiteten hamnar i det kritiska området.
   C och T skall väljas så att
   $$\alpha = P(T(X) \in C) = P(\text{"Förkasta } H_0 \text{ om } H_0 \text{ är sann"})$$

Bra och dåliga hypotestest

Hypotestest på nivån α

H_0 **förkastas** (tillfördel för H_1) med sannolikhet α då (trots att) H_0 är sann.

Kan vi konstruera ett test som **inte använder data**?

Hur beter sig ett **perfekt (idealt)** test?

Styrkefunktion & Felrisker

Styrkefunktion (Kap. 7.6)

Användas för att avgöra hur bra testet skiljer H_0 från H_1.

$$h(\vartheta) = \mathrm{P}(\text{"Förkasta } H_0 \text{ om } \vartheta \text{ är rätt värde"})$$

Typ 1 fel: $\alpha = P(H_0 \text{ förkastas om } H_0 \text{ sann})$
Typ 2 fel: $\beta = P(H_0 \text{ förkastas } \textbf{ej} \text{ om } H_0 \textbf{ ej} \text{ sann})$

Vi ser att $\alpha = h(\vartheta_0)$. Om rätt värde på ϑ är ϑ_1 fås $\beta = 1 - h(\vartheta_1)$.

Naturens okända sanning

		H_0 sann	H_1 sann
Vårt	H_0 förk. ej		β
Beslut	H_0 förkastas	α	

Exempel: Rattonykterhet

Gränsen för rattonykterhet är 0.2 ‰. Antag att mätvärdet vid en mätning, x_i, är

$$X_i = \mu + \varepsilon_i$$

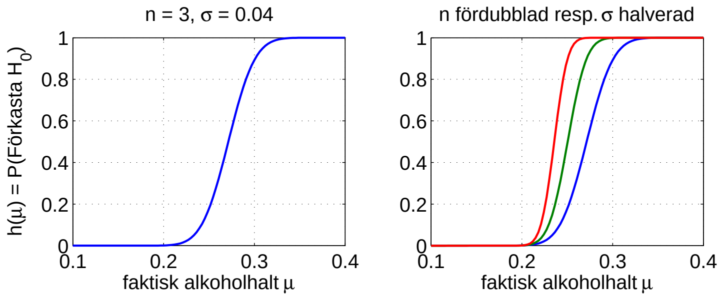
där μ är den sanna alkoholhalten och ε_i är oberoende, $\mathrm{N}(0, 0.04^2)$-fördelade mätfel. För att avgöra om en person är skyldig till rattonykterhet kan man testa hypotesen

$$H_0 : \mu = 0.2 \qquad \text{mot} \qquad H_1 : \mu > 0.2$$

på nivån $\alpha = 0.001$.

1. Ange i ord vad felrisken
$$\alpha = \mathrm{P}(\text{Förkasta } H_0 \text{ om } H_0 \text{ är sann})$$
innebär i hypotestestet ovan.
2. Om man gjort $n = 3$ mätningar och fått medelalkoholhalten till $\bar{x} = 0.24$ ‰, skall man då dömas?
3. Hur högt kan det uppmätta medelvärdet, baserat på tre mätningar, vara utan att man döms?
4. Bestäm styrkefunktionen för testet. Dvs
$$h(\mu) = \mathrm{P}(H_0 \text{ förkastas om } \mu \text{ är det sanna värdet})$$
5. Om den sanna alkoholhalten är 0.25 ‰, vad är då sannolikheten att **inte** dömas.
6. Hur många mätningar behöver göras för att man med högst $\beta = 20\%$ slh skall frikännas om man har 0.25 ‰.

Styrkefunktion för testet av promillehalt ($H_0 : \mu = 0.2$)



Den okända sanningen	Nykter	Olovligt påverkad	
Mätresultat $\mu^* = \bar{x}$		Säkerhetsmarginal	Kritiskt område
Slutsats från test	Frikänns		Döms
	0	0.2	0.27 → μ