Föreläsning 10, Matematisk statistik 7.5 hp för E
Hypotesprövning

Stas Volkov

Konfidensintervall

Ett **konfidensintervall** för en parameter ϑ täcker rätt värde på ϑ med sannolikheten $1-\alpha$.

$1-\alpha$ kallas **konfidensgrad**. Vanliga värden är 0.95, 0.99 och 0.999.

Intervall för normal(approximation)

Om $\vartheta^* \in N(\vartheta, \mathrm{D}(\vartheta^*))$ eller $\vartheta^* \underset{\sim}{\in} N(\vartheta, \mathrm{D}(\vartheta^*))$:
Om $\mathrm{D}(\vartheta^*)$ känd:

$$I_\vartheta = \vartheta^* \pm \lambda_{\alpha/2} \cdot \mathrm{D}(\vartheta^*)$$

Om $\mathrm{D}(\vartheta^*)$ okänd, skattas med medelfelet $d(\vartheta^*)$:

$$I_\vartheta = \vartheta^* \pm t_{\alpha/2}(f) \cdot d(\vartheta^*) \quad \text{om } d(\vartheta^*) \text{ innehåller } \sigma^* = s = \sqrt{\frac{Q}{f}}$$

$$\left(Q = \sum_{i=1}^{n}(x_i - \bar{x})^2,\ f = n-1\right)$$

$$I_\vartheta = \vartheta^* \pm \lambda_{\alpha/2} \cdot d(\vartheta^*) \quad \text{annars}$$

Exempel: Fosforhalt

Man har mätt fosforhalten (μg/l) i en viss sjö $n = 4$ gånger:

$$92.7 \quad 110.9 \quad 101.4 \quad 122.9$$

Vi antar att mätningarna är oberoende observationer av $X_i \in N(\mu, \sigma)$ där μ är den sanna fosforhalten i sjön. Vi vet att $\sigma = 16$ och skattar $\mu^* = \bar{x} = 106.975$.

Om fosforhalten i en sjö överstiger 100 μg/l klassas sjön som *hypertrof (övergödd)*. Kan vi påstå att sjön är övergödd, baserat på dessa fyra mätningar?

Hypotesprövning

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

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

på nivån α; **signifikansnivån** (felrisken) α ges av

$$\alpha = \mathsf{P}(H_0 \text{ förkastas, givet att den är faktiskt sann})$$

"För mycket" beror på osäkerheten i skattningen, $\mathsf{D}(\vartheta^*)$, samt på signifikansnivån α.

Normalt är $\alpha = 0.05, 0.01$ eller 0.001.

Mothypoteser (ekvivalent *Alternativ hypoteser*)

De vanligaste mothypoteserna är

$H_1 : \vartheta \neq \vartheta_0$ då H_0 förkastas om ϑ^* avviker för långt från ϑ_0 både uppåt och nedåt.

$H_1 : \vartheta < \vartheta_0$ då H_0 förkastas om ϑ^* är tillräckligt mycket mindre än ϑ_0.

$H_1 : \vartheta > \vartheta_0$ då H_0 förkastas om ϑ^* är tillräckligt mycket större än ϑ_0.

Olika metoder för att utföra hypotestest

1. **Direktmetoden eller P-värde**
   - Antag att H_0 är sann
   - Räkna ut **P-värdet** $p = \mathsf{P}(\text{Få det vi fått eller "värre"})$
   - Förkasta H_0 om $p < \alpha$
2. **Konfidensmetoden**. 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**
   Förkasta H_0 om testkvantiteten hamnar i det kritiska området. C och T skall väljas så att

$$\alpha = P(T(X) \in C) = P(\text{Förkasta } H_0, \text{ givet } H_0 \text{ är sann})$$

Exempel: Fosforhalt(igen)

Stickprov $x_1, \ldots, x_n$ från $X \in N(\mu,\ \sigma)$ där $n = 4$, $\sigma = 16$ och $\mu^* = \bar{x} = 106.975$.

Hypoteser:

$H_0 : \mu = 100$

$H_1 : \mu > 100$ dvs H_0 förkastas om μ^* är tillräckligt mycket > 100.

- Hur långt från 100 kan man förvänta sig att medelvärdet kan hamna om väntevärdet är 100?
  $\Longrightarrow$ **Kritiskt område**
- Är 106.975 orimligt långt från 100?
  $\Longrightarrow$ **Direktmetoden**
- Är 100 ett rimligt väntevärde om skattningen blev 106.975?
  $\Longrightarrow$ **Konfidensintervall**

Lämplig procedur för lösningar:

- **Modell**: Vad är slumpmässigt och vilken fördelning kan det ha? Sätt upp en lämplig modell.
- **Skattning**: Vilken parameter är vi intresserade av, vad i den är okänt och hur skattar vi det?
- **Egenskaper**: Vad har skattningen för egenskaper?
- **Problemet**: Formulera och lös problemet.
- **Slutsats**: Svara på frågan.

Fosfor:

1. Modell: Vi har $n = 4$ oberoende observationer $x_1, \ldots, x_n$ från $X_i =$ "fosformätning nr. i" $\in N(\mu,\ \sigma)$ där $\sigma = 16$ är känd.

2. Skattning: Vi skattar μ med $\mu^* = \bar{x} = 106.975$.

3. Egenskaper: Vi har att

$$\mu^* = \bar{X} \in N\left(\mu,\ \frac{\sigma}{\sqrt{n}}\right) = N\left(\mu,\ \frac{16}{\sqrt{4}}\right) = N(\mu,\ 8).$$

4. Problemet: Vi vill testa $H_0 : \mu = 100$ mot $H_1 : \mu > 100$ på signifikansnivån (t.ex.) $\alpha = 0.05$.

Om H_0 är sann så gäller att $\mu^* \in N\left(\mu_0,\ \frac{\sigma}{\sqrt{n}}\right) = N(100,\ 8)$.

4. (forts):

- Lösning m.h.a. direktmetoden. P-värdet för testet blir

$$\begin{aligned} p &= P(\text{få det vi fått eller värre om } H_0 \text{ är sann}) \\ &= P(\bar{X} > \bar{x} \mid \mu = \mu_0) = P(\bar{X} > 106.975 \mid \mu = 100) \\ &= 1 - \Phi\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(\frac{106.975 - 100}{8}\right) \\ &\approx 1 - \Phi(0.87) \approx 0.19 \end{aligned}$$

  Eftersom $p = 0.19 \not< \alpha = 0.05$ kan H_0 kan inte förkastas på signifikansnivån $5\,\%$.

- Lösning med hjälp av konfidensintervall: ...
- Lösning med hjälp av kritiskt område: ...

5. Slutsats: Nej, vi kan inte påstår att sjön är övergödd.

Normalfördelning, $\vartheta^* \in \mathsf{N}(\vartheta, \mathsf{D}(\vartheta^*))$, $H_0: \vartheta = \vartheta_0$

Testkvantitet:

$$T = \frac{\vartheta^* - \vartheta_0}{\mathsf{D}(\vartheta^*)}, \qquad \mathsf{D}(\vartheta^*) \text{ känd}$$

$$T = \frac{\vartheta^* - \vartheta_0}{d(\vartheta^*)}, \qquad \mathsf{D}(\vartheta^*) \text{ okänd och } \sigma^* = s = \sqrt{\frac{Q}{f}}$$

Förkasta H_0 om (kritiska områden)

	$H_1: \vartheta < \vartheta_0$	$H_1: \vartheta \neq \vartheta_0$	$H_1: \vartheta > \vartheta_0$
$\mathsf{D}(\vartheta^*)$ känd	$T < -\lambda_\alpha$	$\lvert T\rvert > \lambda_{\alpha/2}$	$T > \lambda_\alpha$
$\mathsf{D}(\vartheta^*)$ okänd	$T < -t_\alpha(f)$	$\lvert T\rvert > t_{\alpha/2}(f)$	$T > t_\alpha(f)$

Jämför kvantiler λ_α eller t_α med konfidensintervallen.

Styrkefunktion & Felrisker

Styrkefunktion

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

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

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

Naturens okända sanning

		H_0 sann	H_1 sann
Vårt beslut	H_0 förk. ej	☺	β
	H_0 förkastas	α	☺

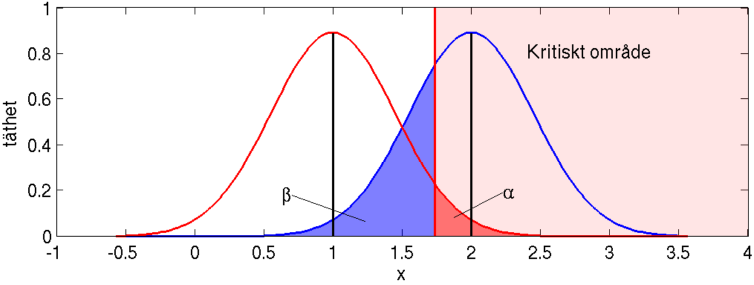
Vi ser att $\alpha = h(\vartheta_0)$. Om rätt värde på ϑ är ϑ_1 fås $\beta = 1 - h(\vartheta_1)$.

Felrisker

Antag $n = 5$ observationer från en $\mathsf{N}(\mu, 1)$ fördelning.
Om vi vill testa (med hjälp av $\mu^* = \frac{1}{5}(x_1 + \cdots + x_5)$)

$$H_0 : \mu = 1$$
$$H_1 : \mu > 1$$

blir fördelningen för μ^* då H_0 är sann (respektive om $\mu = 2$):



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, $N(0, 0.04)$-fördelade mätfel.

För att avgöra om en person är skyldig till rattonykterhet kan man testa hypotesen

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

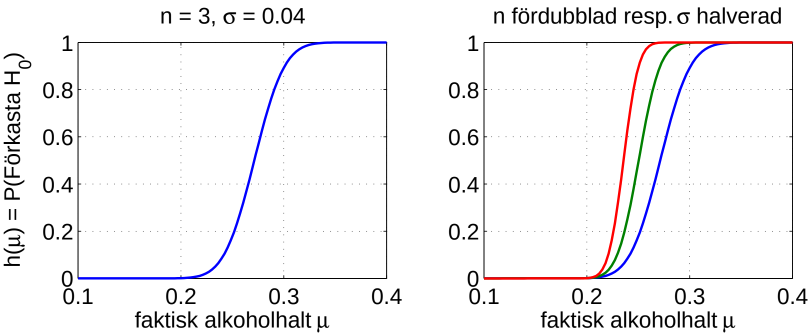
på nivån $\alpha = 0.001$.

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 3 mätningar, vara utan att man döms?
4. Bestäm styrkefunktionen för testet. Dvs. få reda på

$$h(\mu) = \mathrm{P}(H_0 \text{ förkastas om } \mu \text{ är det sanna värdet})$$

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 $\beta = 20\%$ sannolikhet skall frikännas om man har 0.25 ‰. (Dvs bestäm n så att $h(0.25) \geq 1 - \beta$, där $\beta = 0.2$.)

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



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

$0 \qquad 0.2 \qquad 0.27 \qquad \mu$