Institut for Matematik
AARHUS UNIVERSITY

Introduktion til Sandsynlighedsteori og Statistik
Markus Kiderlen
November 9, 2023

UGESEDDEL 12

Forelæsningerne i Uge 10: Vi har færdiggjort Afsnittene 8.2 og 8.3 og vi er startet på Afsnit 8.4 hvor vi er nået til side 478 (def. af ´signifikansniveau α´).

Forelæsningerne i Uge 11:

1. **Forelæsning 21 (14. november):** Vi vil færdiggøre Afsnit 8.4.2 og derefter gennemgå Afsnit 8.4.3 samt starter på Afsnit 8.4.4.

2. **Forelæsning 22 (16. november):** Vi færdiggør Afsnit 8.4.4, gennemgår ret hurtigt Afsnit 8.4.5 og starter derefter på Afsnit 8.5. I forbindelse med Afsnit 8.5 vil vi også vise at estimaterne $\hat{\beta}_0$ og $\hat{\beta}_1$ i boksen på side 502 er maksimum likelihood estimaterne (MLE) (se siderne 511-512 i bogen).

Bemærkninger:

1. **Styrken for et test:** Styrken for et test er $\pi(\theta) = P(W \notin A; \theta)$ for $\theta \in S_1$, dvs. styrken angiver sandsynligheden for at forkaste en falsk nulhypotese. For et givet signifikansniveau vil vi gerne konstruere et test med så stor styrke som muligt, da der så vil være stor sandsynlighed for at forkaste falske nulhypoteser.

2. **P-værdier:** Lad w_1 angive den observerede værdi af vores teststatistik W, hvor $W = \sqrt{n}(\bar{X} - \mu_0)/\sigma$ ved kendt varians og $W = \sqrt{n}(\bar{X} - \mu_0)/S$ ved ukendt varians. I forbindelse med vores gennemgang af P-værdier så vi at disse kan udregnes som følger.

Hvis $X_1, \ldots, X_n$ er en normalfordelt stikprøve så har vi følgende eksakte udtryk for P-værdierne:

	tosided $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	**ensided** $H_0: \mu \leq \mu_0$ (eller $\mu = \mu_0$) $H_1: \mu > \mu_0$	**ensided** $H_0: \mu \geq \mu_0$ (eller $\mu = \mu_0$) $H_1: \mu < \mu_0$
Var. kendt	$2(1 - \Phi(\lvert w_1 \rvert))$	$1 - \Phi(w_1)$	$\Phi(w_1)$
Var. ukendt	$2(1 - F_{T(n-1)}(\lvert w_1 \rvert))$	$1 - F_{T(n-1)}(w_1)$	$F_{T(n-1)}(w_1)$

Hvis $X_1, \ldots, X_n$ er en stikprøve med mange observationer (n er stor, f.eks. $n \geq 50$) så har vi følgende asymptotiske udtryk for P-værdierne:

	tosided	ensided	ensided
	$H_0: \ \mu = \mu_0$ $H_1: \ \mu \neq \mu_0$	$H_0: \ \mu \leq \mu_0$ (eller $\mu = \mu_0$) $H_1: \ \mu > \mu_0$	$H_0: \ \mu \geq \mu_0$ (eller $\mu = \mu_0$) $H_1: \ \mu < \mu_0$
Varians kendt eller ukendt	$2(1 - \Phi(\lvert w_1 \rvert))$	$1 - \Phi(w_1)$	$\Phi(w_1)$

Teoretiske øvelser i Uge 12 (20.-26. november): Del 1.3 og Del 2.3 omhandler lineær regression (Afsnit 8.5) der først gennemgåes til forelæsningen torsdag d. 16. november.

Del 1:

1. Eksamen, Sommer 2019, Opgave 6.
2. Eksamen, Vinter 2018/2019, Opgave 7.
3. Sektion 8.6: Øvelse 21.

Del 2:

1. Eksamen, Vinter 2019/2020, Opgave 5. Bestem også styrken π for testet.
2. Eksamen, Sommer 2018, Opgave 5.
3. Eksamen, Vinter 2020/2021, Opgave 7.

Afleveringsopgave 12: Vi vil gerne teste nulhypotesen at mindst 10% af studerende lider af allergi. Vi har indsamlet data på 225 studerende hvor 21 af dem lider af allergi. Betragt følgende statistiske model for vores data: Lad $X_1, \ldots, X_{225}$ betegne stikprøven hvor $X_i \sim \mathit{Bernoulli}(\theta)$ hvor $\theta \in (0,1)$ er en ukendt parameter. Betragt følgende hypoteser:

$$\begin{aligned} H_0 &: \ \theta \geq 0.1. \\ H_1 &: \ \theta < 0.1. \end{aligned}$$

(a) Argumenter for at ovenstående model svarer til en stikprøve med ukendt middelværdi og varians og argumenter ligeledes for at H_0 svarer til udsagnet at mindst 10% af de studerende lider af allergi.

(b) Lav et asymptotisk hypotesetest med signifikansniveau $\alpha = 0.05$ for nulhypotesen.

(c) Udregn den tilhørende P-værdi.

(Denne opgave er tæt tilknyttet til Exercise 20 i Afsnit 8.6.)

Ugens udfordring: Betragt en stikprøve med én observation $X \sim \mathit{Binomial}(n, p)$ for $n \in \mathbb{N}$ og $p \in (0,1)$. Vi antager at n er en kendt parameter og p er ukendt. Sæt $\theta = 1/p$. Vis at der ikke eksisterer et unbiased estimat $\hat{\Theta}$ for θ.