STK1110 Høsten 2021

P-verdi og noen kommentarer til hypotesetesting

Tilsvarer Avsnitt 9.4 og 9.5 (s. 467-469)

Ingrid Hobæk Haff
Matematisk institutt
Universitetet i Oslo

P-verdi

Eksempel

Eksempel med salve.

- Så langt har har vi bare forholdt oss til en skarp grense for å forkaste H_0, altså enten forkaste eller ikke forkaste.
- Av og til kan en imidlertid være interessert i å vite hvor langt en var fra å forkaste H_0 dersom en ikke forkastet H_0, eller hvor langt er var fra å ikke forkaste, dersom en forkastet H_0.
- **P-verdien** gir informasjon om hvor stor "beviskraft" det er i dataene for å forkaste H_0.
- P-verdien er definert som sannsynligheten for å få en verdi av testobservatoren som **motsier** H_0 (i favør av H_a) **minst like mye** som den observerte verdien, **gitt at** H_0 **er sann**.

P-verdi (forts.)

- Anta at $X_1, \ldots, X_n$ er uif med tetthet/punktsannsynlighet $f(x;\theta)$, og la $V = V(X_1, \ldots, X_n)$ være testobservatoren vi vil bruke til å teste en hypotese angående θ.
- Videre, la v_{obs} være den observerte verdien av V.
- For en test av typen $H_0 : \theta \leq \theta_0$ mot $H_a : \theta > \theta_0$ forkaster vi typisk H_0 for store verdier av V (altså for $V \geq k_1$). Da er P-verdien lik $\mathrm{P}(V \geq v_{obs} | \theta \leq \theta_0)$.
- For en test av typen $H_0 : \theta \geq \theta_0$ mot $H_a : \theta < \theta_0$ forkaster vi typisk H_0 for små verdier av V (altså for $V \leq k_2$). Da er P-verdien lik $\mathrm{P}(V \leq v_{obs} | \theta \leq \theta_0)$.

P-verdi (forts.)

- For en test av typen $H_0 : \theta = \theta_0$ mot $H_a : \theta \neq \theta_0$ forkaster vi typisk H_0 både for små og store verdier av V (altså for $V \leq k_{3L}$ eller $V \geq k_{3U}$).
- Da er det generelt ikke så rett fram å regne ut P-verdien.
- Hvis fordelingen til V er symmetrisk om 0, slik som standard normal- og t-fordeling, er P-verdien gitt ved

$$\begin{aligned} &\mathrm{P}(V \leq -|v_{obs}||\theta \leq \theta_0) + \mathrm{P}(V \geq |v_{obs}||\theta \leq \theta_0) \\ &= 2\mathrm{P}(V \geq |v_{obs}||\theta \leq \theta_0). \end{aligned}$$

- I andre tilfeller trenger vi en alternativ definisjon av P-verdien, nemlig at P-verdien er **det minste signifikansnivået** vi kan bruke hvis vi vil få **forkastet** H_0.

Eksempler

Eksempel

Forsøk med salve.

Eksempel

Eks. 9.8 i boka.

Eksempel

Eks. 9.9 i boka.

Fordelingen til P-verdien

- P-verdien vil variere fra utvalg til utvalg, slik som v_{obs}.
- Det betyr at P-verdien er en stokastisk variabel, akkurat slik som V.
- Anta at vi forkaster H_0 for store verdier av V, og at V er kontinuerlig fordelt.
- Den kumulative fordelingen til V når $\theta = \theta_0$, der θ_0 er nullverdien, er da $F_0(v) = \mathrm{P}(V \leq v | \theta = \theta_0)$.
- La v_{obs} være den observerte verdien av V i en gitt studie. Da er P-verdien for studien gitt ved $\mathrm{P}(V > v_{obs} | \theta = \theta_0) = 1 - F_0(v_{obs})$.
- Det betyr at fordelingen til P-verdien er $1 - F_0(V)$, som er den uniforme fordelingen på $(0, 1)$.
- Dette gjelder også generelt.

Eksempel

Eks. 9.19 i boka.

Statistisk signifikans og fortolkning av P-verdien

- Vi sier at vi har et **statistisk signifikant** resultat dersom P-verdien er mindre enn signifikansnivået, typisk 5%.
- Da forkaster vi H_0, og kan med rimelig grad av sikkerhet anta at H_a er sann.
- P-verdien gir i tillegg informasjon om
  - det er klar forkastning av H_0 eller ikke hvis H_0 er forkastet.
  - en er nær ved å forkaste H_0 eller ikke hvis H_0 ikke er forkastet.

Statistisk signifikans og praktisk relevans

- Merk at det å ikke forkaste H_0 er ikke det samme som å ha vist med rimerlig grad av sikkerhet at H_0 er sann.
- Det kan tenkes at H_0 ikke er sann, men at dataene ikke har nok "beviskraft" til å vise det, f.eks. at utvalget er for lite.
- Videre er et **statistisk signifikant** resultat ikke nødvendigvis det samme som et resultat av **praktisk relevans**.
- Når n er stor, er det stor sannsynlighet for å forkaste H_0, altså få et statistisk signifikant resultat, selv om forskjellen mellom den sanne θ og nullverdien θ_0 er så liten at den ikke er relevant i praksis.
- Fortolkning av P-verdier og betydningen av statistisk signifikans har vært mye diskutert de siste årene, se for eksempel denne artikkelen på forskning.no