

Fortegnstestoppgave

Per og Pål spiller et ballspill. Hver runde består av ett kast fra hver, og rundevinneren er den som kaster lengst. Et spill består av 20 runder. La Y_i betegne lengden på til Pers kast nr. i , og Z_i lengden på Påls kast nr. i .

Per påstår at han er klart bedre enn Pål til å kaste langt.

a) Formuler en hypotesetest for situasjonen ved å spesifisere H_0 og H_1 , velge en passende testobservator, og utlede en beslutningsregel for signifikansnivå $\alpha = 0.1$. Obs: Husk å bruke de stokastiske variablene Y_i og Z_i i uttrykkene dine der det er relevant.

b) Etter 20 runder har Per vunnet 12 og Pål 8. Hva blir konklusjonen av hypotesetesten?

c) Hva blir p -verdien basert på de samme resultatene?

Løsningsforslag fortegnstestoppgave

a) Vi lar $X_i = Y_i - Z_i$, slik at X_i er positiv når Per kaster lengst (og negativ når Pål kaster lengst). La $\tilde{\mu}$ betegne medianen i fordelingen til X_i .

Når Per skal være best, betyr det at X_i må ha en fordeling med median større enn 0 (fordi han må ha flest kast der han har kastet lengst). Dermed får vi følgende hypoteser:

$$H_0 : \tilde{\mu} = 0 \text{ vs. } H_1 : \tilde{\mu} > 0.$$

En naturlig testobservator er antall kast der Per har vært best, altså antall ganger X_i er positiv: $U = \sum_{i=1}^n I(X_i > 0)$. Her er $I(X_i > 0)$ en indikator-funksjon som er 1 når $X_i > 0$ og 0 ellers. Under H_0 er U binomisk fordelt med $n = 20$ og $p = 0.5$, fordi vi under H_0 antar at de to spillerne kaster like godt, og dermed forventer vi at Per kaster lengst like mange ganger som Pål kaster lengst (i det lange løp).

Vi kommer til å ta en beslutning (beholde/forkaste nullhypotesen) om vi observerer et høyt nok antall kast der Per er best. Dermed får vi: "Forkast H_0 om $U \geq k$ " som beslutningsregel, der vi nå må finne verdien til k .

Ved å bruke definisjonen av signifikansnivå kan vi finne k :

$$\begin{aligned} P(\text{forkast } H_0 | H_0 \text{ sann}) &= P(U \geq k | H_0) \\ &= 1 - P(U < k | H_0) \\ &= 1 - P(U \leq k - 1 | H_0) \leq \alpha \\ 1 - \alpha &\leq P(U \leq k - 1 | H_0) \\ P(U \leq k - 1 | H_0) &\geq 1 - \alpha = 0.9 \end{aligned}$$

Husk: Under H_0 er $U \sim \text{Bin}(n = 20, p = 0.5)$. Vi slår derfor opp i tabellen for binomisk fordeling med $n = 20$, kolonne for $p = 0.5$, og finner at:

$P(U \leq 12) = 0.868$ og $P(U \leq 13) = 0.942$. Dermed får vi at $P(U \leq 13 | H_0) \geq 0.9$, og $k - 1 = 13 \implies k = 14$. Beslutningsregelen blir altså at vi forkaster nullhypotesen om at Per og Pål er like gode hvis Per kaster lengst i minst 14 av de 20 rundene.

b) Per vant (kastet lengst) i 12 av de 20 rundene, og vi har ikke grunnlag for å forkaste nullhypotesen.

c) p -verdi er definert som sannsynligheten for å observere det vi har observert eller noe mer ekstremt i retning alternativhypotesen, gitt at H_0 er sann. I vårt

tilfelle blir det: $P(U \geq 12|H_0)$.

$$\begin{aligned}P(U \geq 12|H_0) &= P(U \geq 12|U \sim \text{Bin}(n = 20, p = 0.5)) \\&= 1 - P(U < 12|U \sim \text{Bin}(n = 20, p = 0.5)) \\&= 1 - P(U \leq 11|U \sim \text{Bin}(n = 20, p = 0.5)) \\&= 1 - 0.748 = 0.252\end{aligned}$$