Oppgave 1

Du trekker fem kort uten tilbakelegging fra en vanlig kortstokk med 52 kort. Det er 13 kort av hver av de fire kortfargene spar, kløver, ruter og hjerte. I denne oppgava kan du velge om du vil finne teoretiske eller simulerte sannsynligheter.

a)

Hva er sannsynligheten for å trekke fem spar?

b)

Hva er sannsynligheten for at alle de fem korta har samme kortfarge?

c)

Hva er sannsynligheten for å trekke tre spar og to ruter?

Oppgave 2

Av alle som tar offentlig transport i Oslo en dag kaller vi andelen som ikke har kjøpt billett for *p*. Ruter kontrollerer *n* reisende og kaller antallet personer uten billett for *X*. Vi forutsetter at de kontrollerte er et tilfeldig utvalg av de reisende.

a)

Hvilken fordeling har *X*?

b)

Vi forutsetter nå at det er en kostnad $c(n) = kn^2$ ved kontroll av *n* personer, der *k* er en konstant. Hvis en person ikke har kjøpt billett må han betale ei bot *b*. Den daglige profitten, π, for billettkontrollørene til Ruter ved kontroll av *n* personer er derfor:

$$\pi = bX - c(n) = bX - kn^2$$

Finn forventa daglig profitt ved billettkontroller hvis en kontrollerer *n* personer. Vi kaller nivået av *n* som maksimerer den forventa profitten for n^* . Vis at n^* er lik:

$$n^* = p * \frac{b}{2k}$$

Du kan ta for gitt at den forventa profittfunksjonen er konkav.

c)

Siden *p* er ukjent er det optimale valget av *n* ukjent, og du må estimere n^*. Ruter engasjerer deg til å finne ut hvor mange de må kontrollere hver dag for å maksimere profitten ved billettkontroller. Du kontrollerer *n* tilfeldige reisende en dag og foreslår følgende estimator:

$$\widehat{n^*} = \frac{X}{n} * \frac{b}{2k}$$

Er denne estimatoren forventningsrett? Hva er variansen til estimatoren? Finn estimatet for n^* når $k = 0.1$, $n = 500$, $b = 1500$ og du observerer $X = 60$.

d)

Konstruer et 99% konfidensintervall for n^* basert på estimatet ditt i c). Vær klar på hvilke forutsetninger du bruker.

e)

Maksimal testkapasitet for Ruter er 1000 kontrollerte personer hver dag. De ønsker derfor å vite om n^* er mindre enn 1000. Konstruer en test for å avgjøre om n^* er mindre enn 1000 basert på estimatet ditt i c). Sett opp passende hypoteser og signifikansnivå og gjennomfør testen. [Hint: Hvis $n^* = 1000$ impliserer dette at $p = 0.13$. Du trenger ikke å vise dette.]

f)

Hva er p-verdien til hypotesetesten i e)?

Oppgave 3

Du fisker ved et vann. Hvor mange fisk du får i løpet av en fisketur på fire timer er Poisson-fordelt med parameter $\lambda = 4$.

a)

Diskuter hva det innebærer å forutsette at hvor mange fisk du får er Poisson-fordelt.

b)

Vi later nå som om du ikke vet hvilken verdi λ har, og at du må estimere denne. Du reiser på $n = 50$ fisketurer, alle på fire timer hver, og kaller den totale mengden fisk du får for X. Du foreslår å bruke estimatoren:

$$\hat{\lambda} = \frac{X}{n}$$

Forklar kort hvorfor du kan bruke tilnærminga:

$$\hat{\lambda} \sim N(4, 4/50)$$

c)

Bruk tilnærminga fra b) til å finne den tilnærma sannsynligheten $P(\hat{\lambda} \leq 3.5)$. Bruk R til å finne den eksakte sannsynligheten $P(\hat{\lambda} \leq 3.5)$ og sammenlign med tilnærminga.

Ta for deg følgende hypotesetest med signifikansnivå α:

$$H_0: \lambda = \lambda_0, \quad H_1: \lambda \neq \lambda_0$$

d)

Forklar hva en Type I-feil er. Hva er sammenhengen mellom sannsynligheten for Type I-feil og signifikansnivået α?

e)

Forklar hva en Type II-feil er. Hva er sammenhengen mellom sannsynligheten for Type II-feil og signifikansnivået α?

f)

Du skal nå simulere teststyrke i R. Du skal gjennomføre hypotesetesten ovenfor med $\lambda_0 = 3.5$ og signifikansnivå $\alpha = 0.01$. Siden den faktiske $\lambda = 4$, veit du at nullhypotesen ikke er riktig. Trekk tilfeldige verdier av 50 fisketurer og gjennomfør hypotesetesten i R. Repeter dette 1000 ganger. Finn hvor stor del av gangene du forkaster den gale nullhypotesen. Tolk resultatet.