TMA4245 Statistikk
Vår 2019

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Innlevering 3

Dette er tredje av tre innleveringer i blokk 1. Denne øvingen skal oppsummere pensum forelest de seks første forelesningsukene. Spesielt er det i denne innleveringen fokus på viktige diskrete og kontinuerlige sannsynlighetsfordelinger. For å få godkjent innleveringen kreves det at minimum 40% av svarene er riktige, og at besvarelsen reflekterer at man har gjort et ordentlig forsøk på minimum 75% av oppgavene. Alle oppgaver teller like mye.

Oppgave 1

I semifinalen mellom Frankrike og Tyskland i et verdensmesterskap i fotball er resultatet uavgjort etter ekstraomganger.

Vi definerer en runde i straffesparkkonkuransen til å være at hvert lag skyter en straffe hver.

I første del av straffesparkkonkurransen er det 5 runder, 5 straffespark fra hvert lag, og laget som skårer flest ganger vinner.

Dersom lagene skårer like mange ganger i første del, går man over til andre del av straffesparkkonkurransen. Nå spilles det en og en runde: Hvert lag har en straffe hver. Dersom det ene laget skårer og det andre bommer, har vi en vinner. Ellers spilles en ny runde (hvert lag får en straffe hver) inntil vi har en vinner.

Anta at de tyske spillerene har sannsynlighet $p_T = 0.80$ for å skåre på straffe, at de franske spillerene har sannsynlighet $p_F = 0.70$ for å skåre, og at utfallene av straffesparkene er uavhengige av hverandre.

a) La X_T og X_F være antall mål henholdsvis Tyskland og Frankrike skårer i første del av straffesparkkonkurransen. Hvilken fordeling har X_T og X_F? Begrunn svaret!

Hva er sannsynligheten for at stillingen blir 5-5 etter første del?

Hva er sannsynligheten for at stillingen blir 3-3 etter første del?

La L_1 være hendelsen at lagene står likt etter første del. Sett opp et uttrykk for hvordan man kan finne sannsynligheten $P(L_1)$ for at lagene står likt etter første del.

b) Anta i dette punktet at vi er kommet til del 2 av straffesparkkonkuransen. La D_2 være hendelsen at konkurransen er i del 2, og V_i betegne hendelsen at vi har en vinner i runde i, $i = 1, \ldots$. La videre V_T være hendelsen at Tyskland vinner.

Hva er sannsynligheten for at vi har en vinner etter første runde i del to, $P(V_1|D_2)$?

Gitt at vi har en vinner etter første runde i del 2, hva er sannsynligheten $P(V_T|V_1, D_2)$ for at dette er Tyskland?

c) La X være antall runder i del 2 til og med runden det blir kåret en vinner i.

Hvilken fordeling har X? Begrunn svaret og oppgi parameter/re.

Hva er forventningsverdi og varians i antall runder (i del 2) til det blir kåret en vinner?

Oppgave 2

En forhandler kjøper inn et stort parti kulepenner og selger det videre til kunder. Han får klage på alle kulepenner som ikke virker. Anta først at antallet klager, X, er poissonfordelt med parameter $\lambda > 0$:

$$p(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad ; \ x = 0, 1, 2, \ldots$$

a) Anta, kun i dette punktet, at $\lambda = 7$. Finn sannsynligheten for at forhandleren får en eller flere klager. Finn også sannsynligheten for at forhandleren får færre enn tre klager, gitt at han får en eller flere klager.

Leverandøren av kulepenner har tre fabrikker, benevnt A, B og C, som produserer med ulik kvalitet. Forhandleren mottar varepartiet fra en av fabrikkene men han vet ikke fra hvilken. Dersom varepartiet kommer fra fabrikk A vil antall klager være poissonfordelt med parameter $\lambda_A = 5$, hvis partiet er fra fabrikk B og C er klagefordelingsparameteren henholdsvis $\lambda_B = 15$ og $\lambda_C = 20$. Forhandleren antar i utgangspunktet at det er sannsynlighet $p_A = 0.5$ for at varepartiet kommer fra fabrikk A og sannsynlighet $p_B = p_C = 0.25$ for hver av fabrikkene B og C.

b) Utled et uttrykk for den marginale sannsynlighetsfordelingen for antall klager forhandleren kan regne med å få.

Oppgave 3

En entreprenør har leid inn et transportfirma til å transportere masse bort fra en byggeplass. Det er to mulige veivalg fra byggeplassen til stedet der massen skal deponeres, gjennom eller utenom bykjernen. Av helse, miljø og sikkerhetshensyn velger entreprenøren at massen skal transporteres utenom bykjernen.

Når oppdraget er ferdig har transportfirmaet kjørt 1000 turer, og transportfirmaet informerer entreprenøren om at av de 1000 turene har 5 blitt kjørt gjennom bykjernen og 995 er blitt kjørt utenom bykjernen.

I løpet av transportperioden har entreprenøren ved 5 tilfeldig valgte transporter sjekket om transporten har skjedd gjennom bykjernen. La X være antall ganger, av de 5 transportene som ble sjekket, entreprenøren finner at massen er blitt transportert gjennom bykjernen.

Hvilken fordeling har X? Begrunn svaret.

Hvilken verdi av X har høyest punktsannsynlighet?

Nå viser det seg at entreprenøren fant at i 5 av de 5 tilfellene han sjekket så ble massen transportert gjennom bykjernen. Hva er sannsynligheten for dette, dvs. $P(X = 5)$?

Oppgave 4

En bestemt målemetode for bestemmelse av pH-verdien i en løsning gir måleresultater som antas å være uavhengige og normalfordelte, med forventningsverdi μ lik virkelig pH og varians $\sigma^2 = 0.060^2$. La $X_1, \ldots, X_n$ være uavhengige målinger av pH i en bestemt løsning.

Anta at den virkelige pH-verdien i en løsning er 6.8.

Hva er sannsynligheten for at en bestemt måling gir et resultat som er under 6.74?

Hva er sannsynligheten for at en bestemt måling gir et resultat som er mellom 6.74 og 6.86?

Hva er sannsynligheten for at en bestemt måling, X, gir et resultat som avviker mer enn 0.06 fra μ, dvs bestem $P(|X-\mu| > 0.06)$?

Oppgave 5

En fabrikk produserer kabel og en gang i blant oppstår det feil på den produserte kabelen. La Z betegne lengden (i kilometer) på kabelen mellom to etterfølgende feil. Vi skal anta at feilene oppstår uavhengig av hverandre, dvs. at påfølgende observasjoner av Z langs kabelen, $Z_1, Z_2, Z_3, \ldots$, er uavhengige stokastiske variable.

Av erfaring vet en at lengden mellom to etterfølgende feil er eksponensialfordelt med parameter λ, dvs. Z har sannsynlighetstetthet

$$f(z;\lambda) = \begin{cases} \lambda e^{-\lambda z} & \text{for } z > 0, \\ 0 & \text{ellers} \end{cases}$$

og kumulativ fordelingsfunksjon

$$F(z;\lambda) = \begin{cases} 1 - e^{-\lambda z} & \text{for } z > 0, \\ 0 & \text{ellers.} \end{cases}$$

a) Anta i dette punktet at $\lambda = 0.05$.

Hva er sannsynligheten for at lengden mellom to etterfølgende feil er mer enn 10 kilometer?

Dersom man har observert at de første 10 kilometrene er feilfrie, hva er da sannsynligheten for at også de neste 10 kilometrene er feilfrie?

I resten av oppgaven skal vi anta at λ er en ukjent parameter. Ved hjelp av fabrikkens opptegnelser over tidligere feil på kabelen ønsker vi å estimere λ. Men det viser seg dessverre at fabrikken ikke har notert nøyaktig lengde på kabelen mellom hver feil, i stedet er det kun notert antall hele kilometer, M, med kabel mellom hver feil. Dvs, dersom $Z < 1.0$ har fabrikken notert seg $M = 0$, dersom $1.0 \leq Z < 2.0$ har fabrikken notert seg $M = 1$, dersom $2.0 \leq Z < 3.0$ har fabrikken notert seg $M = 2$, osv.

b) Vis at punktsannsynligheten for M blir

$$\mathrm{P}(M = m) = (1 - e^{-\lambda})e^{-\lambda m} \quad \text{for} \quad m = 0, 1, 2, \ldots.$$

Fasit

1. a) 0.055, 0.063 b) 0.380, 0.632 c) $E[X] = 2.632$

2. **a**) 0.9991, 0.0288

3. $P(X = 5) = 1.21 \cdot 10^{-13}$

4. $0.159, 0.682, 0.318$

5. **a**) $0.607, 0.607$