Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

TMA4240 Statistikk
Høst 2017

Anbefalt øving 6

Denne anbefalte øvinga tar utgangspunkt i pensum i sjette uke med forelesninger. Oppgavene handler om kontinuerlige sannsynlighetsfordelinger, særlig normal- og eksponentialfordelingene.

Oppgave 1

I denne oppgaven kan du bruke uten å vise det at

$$\int_0^\infty x^r e^{-ax} dx = \frac{r!}{a^{r+1}} \text{ når } a > 0 \text{ og } r \text{ er et heltall} \geq 0$$

Vi betrakter ankomst- og oppholdstider for et bestemt lokaltog på en jernbanestasjon. Toget skal etter rutetabellen ankomme hver hverdag klokka 8:00, men kommer alltid etter dette tidspunktet.

La X (minutter) betegne togets forsinkelse på en tilfeldig valgt hverdag. Vi antar at X er en stokastisk variabel med sannsynlighetstetthet

$$g(x) = \begin{cases} kxe^{-2x} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

der $k > 0$ er en konstant.

a) Vis at $k = 4$.

Hva er den forventede forsinkelse for toget?

Vis at sannsynligheten for at toget er mer enn 2 minutter forsinket er tilnærmet lik 0.09.

b) La V være antall ganger i løpet av en måned (= 22 hverdager) at toget er mer enn 2 minutter forsinket. Foreslå en sannsynlighetsfordeling for V og sett opp de forutsetninger som ligger til grunn for denne.

Hva er sannsynligheten for at toget er mer enn 2 minutter forsinket minst 2 ganger i løpet av en måned (= 22 hverdager)?

Hva er (tilnærmet) sannsynligheten for at toget er mer enn 2 minutter forsinket mer enn 30 ganger i løpet av 220 hverdager?

La Y (minutter) være den tiden toget står på stasjonen. Oppholdstiden Y vil være influert av forsinkelsen, og vi antar at den betingede sannsynlighetstetthet $f(y|x)$ for Y, gitt at forsinkelsen X er lik x (> 0), er gitt ved

$$f(y|x) = \begin{cases} (x/2)\, e^{-xy/2} & \text{for } y > 0 \\ 0 & \text{for } y \leq 0 \end{cases}$$

c) Hvilken fordeling har oppholdstiden Y når det er gitt at forsinkelsen er 2 minutter?

Hva er forventet oppholdstid når forsinkelsen er 2 minutter?

Sett opp simultantettheten $f(x, y)$ for X og Y.

Finn sannsynlighetstettheten $h(y)$ for oppholdstiden Y.

Oppgave 2

En pakke med egg inneholder 6 egg. Hvert egg antas uavhengig og normalfordelt med forventningsverdi 70 gram og varians 16 gram^2.

a) Hva er sannsynligheten for at et tilfeldig valgt egg veier mer enn 60 gram?

La Y være vekten av en tilfeldig valgt pakke med egg. Hva er fordelingen til Y ? (Begrunn svaret).

b) Produsenten ønsker å angi en garantert minstevekt av en pakke med egg, slik at det er 95% sannsynlighet for at en tilfeldig valgt eske tilfredstiller garantien. Hva blir den garanterte minstevekt?

c) Bruk Matlab til å lage en figur som illustrerer situasjonen i **b)**. Plott sannsynlighetstettheten til pakkevekten, marker den garanterte minstevekten, og fargelegg området under grafen og til venstre for minstevekten.

Oppgave 3

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 4 Anta at levetiden til en bestemt type elektroniske komponenter er eksponensialfordelt med forventningsverdi lik $1/\lambda$. Det finnes mange produsenter av denne typen elektroniske komponenter og kvaliteten på produktet varierer fra produsent til produsent. Dvs. de forskjellige produsentene har forskjellig parameterverdi λ og verdien på λ beskriver dermed gjennomgående kvalitet på komponenter fra den enkelte produsent. Anta videre at dersom en tilfeldig velger en produsent så kan en betrakte tilhørende λ som en kontinuerlig fordelt stokastisk variabel som er eksponensialfordelt med forventningsverdi $1/\theta$.

Anta at en kunde, som skal kjøpe en elektronisk komponent, går frem på følgende måte: Først velger han tilfeldig en produsent og deretter går han og kjøper en komponent produsert av denne produsenten. La T betegne levetiden for den komponenten kunden kjøper.

a) Finn sannsynlighetsfordelingen for T.

Sannsynlighetsfordelingen du fant i **a)** kan verifiseres ved hjelp av simulering i Matlab. La $\theta = 8$ i resten av oppgaven.

b) Anta at en kunde skal kjøpe 100 komponenter. Lag et program som simulerer levetiden til hver av de 100 komponentene kunden kjøper. Det vil si:
1) La kunden kjøpe en komponent fra en tilfeldig produsent. Siden parameterverdien λ er eksponensialfordelt må den simuleres fra en eksponensialfordeling.
2) Simuler levetiden til denne spesifikke komponenten gitt parameterverdien λ som ble funnet i steg 1).
3) Gjenta dette 100 ganger og ta vare på levetiden som oppnås for hver iterasjon.
For å generere en tilfeldig verdi fra en eksponensialfordeling kan du bruke funksjonen `exprnd()` i Matlab.

c) Lag et histogram av de simulerte levetidene fra oppgave **b)**. Plott den teoretiske sannsynlighetsfordelingen T funnet i **a)** i samme figur og kommenter svaret.

Oppgave 5

Vis at eksponensialfordelingen er "glemsk" (har ingen hukommelse), dvs. vis at

$$P(X \geq t + s | X > s) = P(X \geq t) \quad \forall\ s, t > 0.$$

Vis tilsvarende for geometrisk fordeling, dvs. vis at

$$P(X \geq t + s | X > s) = P(X \geq t) \quad \forall\ s, t > 0.$$

Fasit

2. **a**) P($X > 60$)=0.994, Garantert minstevekt: 403.9g

3. $0.159, 0.682, 0.318$

4. $f(t) = \theta/(t + \theta)^2$