TMA4240/TMA4245 Statistikk
Ordningsvariabler og ekstremvariabler

1 Ordningsvariabler

La $X_1, X_2, \ldots, X_n$ være uavhengige, identisk fordelte stokastiske variabler med fordeling $f_X(x)$ og kumulativ fordeling $F_X(x) = P(X \leq x)$.

Vi ordner X_i-ene etter størrelse og betegner dem $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$, hvor

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(k)} \leq \ldots \leq X_{(n)}.$$

Vi kaller $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ ordningsvariabler.

Spesielt definerer vi ekstremvariablene

$$\begin{aligned} X_{(1)} &= \min(X_1, X_2, \ldots, X_n) \\ X_{(n)} &= \max(X_1, X_2, \ldots, X_n). \end{aligned}$$

Medianen er definert ved:

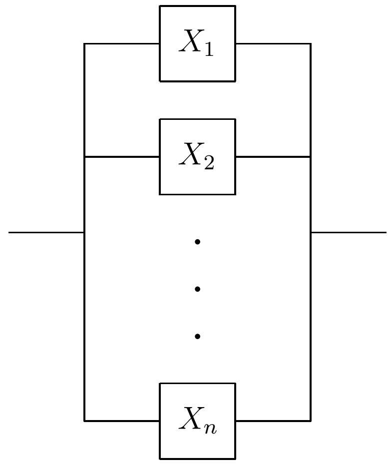
$$\tilde{X} = \begin{cases} X_{((n+1)/2)} & \text{hvis } n \text{ er oddetall} \\ \frac{1}{2}(X_{(n/2)} + X_{(n/2+1)}) & \text{hvis } n \text{ er partall} \end{cases}$$

Variasjonsbredden ("range" på engelsk) er definert ved: $X_{(n)} - X_{(1)}$.

2 Maksimum

2.1 Eksempel: Parallellsystem og maksimum

Et eksempel på en situasjon der fordelingen til ekstremvariabler er interessant, er levetiden til et system sammensatt av flere delkomponenter. La oss se på levetiden til et system sammensatt av komponenter med uavhengige levetider. I systemer eller delsystemer der det stilles høye krav til at systemet skal virke kobles ofte komponenter i parallell:



Systemet virker så lenge minst en av komponentene virker, og levetiden til systemet blir da $V = \max(X_1, X_2, \ldots, X_n)$, der X_i er levetiden til komponent i.

2.2 Fordelingen til maksimum

Fordelingen til $V = \max(X_1, X_2, \ldots, X_n)$ kan finnes ved å benytte at den største av X_i-ene er mindre enn eller lik v hvis og bare hvis alle X_i-ene er mindre enn eller lik v:

$$\begin{aligned} F_V(v) = P(V \leq v) &= P(\max(X_1, X_2, \ldots, X_n) \leq v) \\ &= P((X_1 \leq v) \cap (X_2 \leq v) \cap \ldots \cap (X_n \leq v)) \\ &\overset{uavh.}{=} P(X_1 \leq v) \cdot P(X_2 \leq v) \cdots P(X_n \leq v) \\ &= [F_X(v)]^n \end{aligned}$$

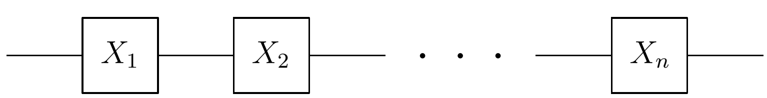
Hvis X er kontinuerlig fordelt blir:

$$f_V(v) = \frac{d}{dv} F_V(v) = n[F_X(v)]^{n-1} f_X(v)$$

3 Minimum

3.1 Eksempel: Seriesystem og minimum

La oss fortsatt se på eksemplet med et system av komponenter med uavhengige levetider, men anta nå at systemet virker hvis og bare hvis samtlige komponenter virker. Dette kan illustreres ved en seriekobling:



La X_i være levetiden til komponent nr i. Systemet funksjonerer frem til første komponent svikter. Da er levetiden til systemet $U = \min(X_1, X_2, \dots, X_n)$.

3.2 Fordelingen til minimum

Fordelingen til $U = \min(X_1, X_2, \dots, X_n)$ blir på tilsvarende måte som for maksimum:

$$\begin{aligned}
F_U(u) = P(U \le u) &= P(\min(X_1, X_2, \dots, X_n) \le u) \\
&= 1 - P(\min(X_1, X_2, \dots, X_n) > u) \\
&= 1 - P((X_1 > u) \cap (X_2 > u) \cap \dots \cap (X_n > u)) \\
&\overset{uavh.}{=} 1 - P(X_1 > u) \cdot P(X_2 > u) \cdots P(X_n > u)) \\
&= 1 - [1 - F_X(u)]^n
\end{aligned}$$

Hvis X er kontinuerlig fordelt blir:

$$f_U(u) = \frac{d}{du} F_U(u) = n[1 - F_X(u)]^{n-1} f_X(u)$$

3.3 Eksponensialfordeling og minimum:

La $X_1, X_2, \dots, X_n$ være uavhengige, eksponensialfordelte levetider med forventning β. Dvs $f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$ for $x \ge 0$, $\beta > 0$. Fordelingen til $U = \min(X_1, X_2, \dots, X_n)$ blir:

$$\begin{aligned}
F_U(u) = P(U \le u) &= P(\min(X_1, X_2, \dots, X_n) \le u) \\
&= 1 - [1 - F_X(u)]^n \\
&= 1 - [1 - \int_0^u f_X(x) dx]^n \\
&= 1 - [1 - \int_0^u \frac{1}{\beta} e^{-\frac{x}{\beta}} dx]^n \\
&= 1 - [1 - (1 - e^{-\frac{u}{\beta}})]^n \\
&= 1 - e^{-\frac{nu}{\beta}} \qquad\qquad \text{for } x > 0
\end{aligned}$$

Altså er levetiden til seriesystemet eksponensialfordelt med forventning $\frac{\beta}{n}$. I oppgave 2 (til slutt i notatet), skal et tilsvarende resultat vises for en annen fordeling, kalt Weibull-fordelingen.

4 $k$te ordningsvariabel

Fordelingen til $X_{(k)}$ blir

$$F_{X_{(k)}}(x) = P(k \text{ eller flere } X_i\text{-er er} \leq x) = \sum_{j=k}^{n} \binom{n}{j} [F_X(x)]^j [1 - F_X(x)]^{n-j}$$

fordi antallet $X_i \leq x$ er binomisk fordelt med parametre n og $F_X(x)$.

Sannsynlighetstettheten finnes ved å derivere m.h.p. x og etter noe mellomregning kan den skrives:

$$f_{X_{(k)}}(x) = n\binom{n-1}{k-1}[F_X(x)]^{k-1}[1 - F_X(x)]^{n-k} f_X(x)$$

5 Oppgaver

Oppgave 1. Betrakt et parallellsystem av 2 uavhengige komponenter. Levetiden til hver komponent er eksponensialfordelt med parameter λ, dvs

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{ellers} \end{cases}$$

La V være levetiden til systemet. Finn fordelingen til V, samt $\mathrm{E}(V)$.

Oppgave 2. Betrakt et seriesystem sammensatt av n komponenter. Levetiden til hver komponent følger en fordeling med kumulativ fordelingsfunksjonen

$$F_X(x) = \begin{cases} 1 - e^{-(\lambda x)^\alpha} & \text{for } x \geq 0 \\ 0 & \text{ellers} \end{cases}$$

hvor $\lambda > 0$ og $\alpha > 0$. Dette kalles en Weibull-fordeling med skalaparameter λ og formparameter α. (Parametriseringen er litt anderledes enn i læreboka.)

Finn den kumulative fordelingsfunksjonen for levetiden til systemet. Hva kalles denne sannsynlighetsfordelingen?