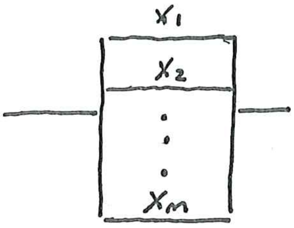
Fordelinga til $\max(X_1, X_2, \dots, X_m)$

La oss sjå på eit parallell system samansatt av komponentar med uavhengige og identisk fordelte levetider $X_1, X_2, \dots, X_m$. $P(X_i \leq x) = F(x)$, $i = 1, 2, \dots, m$



La $V = \max(X_1, X_2, \dots, X_m)$

$$P(V \leq v) = P(X_1 \leq v, X_2 \leq v, \dots, X_m \leq v)$$

$$= P(X_1 \leq v) \cdot P(X_2 \leq v) \cdot \dots \cdot P(X_m \leq v) = F(v)^m$$

$$f_V(v) = m(F(v))^{m-1} f(v)$$

Fordelinga til den k-te ordningsvariabelen

La alle X_i, $i = 1, 2, \dots, m$ vere uavhengige og identisk fordelte

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(k)} \leq \dots \leq X_{(m)}$$

$$P(X_{(k)} \leq x) = F_{X_{(k)}}(x) = P(k \text{ eller fleire } X_i\text{-ar er } \leq x)$$

$$= \sum_{j=k}^{m} \binom{m}{j} F(x)^j (1 - F(x))^{m-j}$$

Prov. La Y vere talet på $X_i \leq x$ Y er $B(m, F(x))$

$$F_{X_{(k)}}(x) = P(Y \geq k)$$

8. Utvalsfordelingar og beskriving av datasett

Definisjonar

8.1. Ein populasjon består av alle observasjonar verdiar av interesse for det fenomenet vi studerer.

8.2. Eit utval er ei delmengde av populasjonen

8.3. Dei tilfeldige stokastiske variablane $X_1, \dots, X_n$ seiast å vere eit tilfeldig utval dersom

1. Dei har same sannsynstettleik $f(x)$
2. $f(x_1, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n)$ ɔ: dei er uavhengige.

Eks.

La X_i vere talet på feil på glas nummer i, $i = 1, 2, \dots, 30$

$X_i \sim \text{Poisson}(\lambda t)$ ɔ: $P(X=x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$. Set $t=1$ s.a.

$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$. $E[X] = \lambda$

X_i : 01110 00201 01000 00102 00312 00100

$\hat{\lambda} = \frac{\sum_{i=1}^{30} x_i}{30} = 0,57$, $s^2 = \sum_{i=1}^{30} \frac{(x_i - \bar{x})^2}{30-1} = 0,67$

NB! I Poissonfordelinga er $E[X] = Var[X]$.

Ordna etter storleik får vi:

00000 00000 00000 00011 11111 1 222 3

Utvals median : 0

Modal verdi : 0 (Den som skjer oftast

Def. 8.4. Ein statistikk er ein funksjon av tilfeldige variable som utgjer eit tilfeldig utval.

Eks. $X_1, \dots, X_n$ stokastiske variable i eit utval.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

8.5. Sannsynsfordelinga til ein statistikk blir kalla ei utvalsfordeling.

Eks.

La $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \dots, n$ vere eit tilfeldig utval.

$$E[\bar{X}] = E\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \sum_{i=1}^{n} \frac{E[X_i]}{n} = \frac{n\mu}{n} = \mu$$

$$Var[\bar{X}] = Var\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} \sum_{i=1}^{n} Var[X_i] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$X_i \sim N(\mu, \sigma^2)$ og uavh. $i = 1, 2, \dots, n \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Kva med $E\left[\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}\right] = E[S^2]$

Vi får
$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} [(X_i - \mu) - (\bar{X} - \mu)]^2$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - 2(\bar{X} - \mu) \underset{\underset{n(\bar{X}-\mu)}{\parallel}}{\sum_{i=1}^{n} (X_i - \mu)} + n(\bar{X} - \mu)^2$$

$$= \sum_{i=1}^{n} (X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2 = \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2$$

s.a. $E\left[\sum_{i=1}^{n} (X_i - \bar{X})^2\right] = \sum_{i=1}^{n} E[(X_i - \mu)^2] - n E[(\bar{X} - \mu)^2]$

$$= n\sigma^2 - n \cdot \frac{\sigma^2}{n} = (n-1)\sigma^2$$

s.a. $E[S^2] = E\left[\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}\right] = \frac{(n-1)\sigma^2}{n-1} = \sigma^2$

Sentralgrense teoremet

Har sett : $X \sim B(n,p)$ der $\begin{cases} np \geq 5 \\ n(1-p) \geq 5 \end{cases}$ så er

$$\frac{X-np}{\sqrt{np(1-p)}} \approx N(0,1)$$

$X = \sum_{i=1}^{n} Y_i$ der $Y_i = \begin{cases} 1 & \text{dersom A skjer i i-te forsøk} \\ 0, & \text{elles.} \end{cases}$

$$\frac{X-np}{\sqrt{np(1-p)}} = \frac{\frac{X}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\bar{Y} - E[Y]}{\frac{\sigma_Y}{\sqrt{n}}} \Longrightarrow \sigma_{\bar{Y}}$$

Generelt Teorem 8.2

La $X_1, \ldots, X_n$ vere eit tilfeldig utval. $E[X_i] = \mu$

$Var[X_i] = \sigma^2$, $i = 1, 2, \ldots, n$

Då er $\lim_{n \to \infty} \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$

Normalt nok at $n \geq 30$