

Sannsynlighetsfordelinger:

Binomisk Fordeling:

- 1. n uavhengige delforsøk
- 2. Suksess eller ikke
- 3. $P(A)=p$ i alle forsøk
 - $X =$ Antall ganger A inntreffer på n forsøk.
 - $X \sim binom(n, p)$
 - $f(x) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}, \ x = 0, 1, 2, \dots, n$
 - $P(X \leq x) = \sum_{k=0}^x P(X = k)$
 - $E(X) = np \quad Var(X) = np(1 - p)$

Hypergeometrisk:

- 1. Populasjon med N elementer.
- 2. k av disse regnes som ”Suksess”, $N - k$ som fiasko
- 3. Trekker n elementer uten tilbakelegging
 - X , antallet suksesser.
 - $f(x) = \frac{\binom{k}{x} \cdot \binom{N-k}{n-x}}{\binom{N}{n}}$
 - $E(X) = np \quad Var(X) = np(1 - p) \frac{N-n}{N-1}, \ p = k/N$

Negativ-Binomisk:

- X er antall forsøk en må gjøre for at en hendelse A skal inntreffe k ganger
- $f(x) = \binom{x-1}{k-1} \cdot p^k (1 - p)^{x-k}, \ x = k, k + 1, k + 2, \dots$
 - $E(X) = k/p \quad Var(X) = k \cdot \frac{1-p}{p^2}$

Geometrisk:

- X er antall forsøk en må gjøre for at hendelsen A inntreffer første gang.
- $E(X) = 1/p \quad Var(X) = \frac{1-p}{p^2}$
- Geometrisk fordeling er minneløs!

Poisson:

- Antall forekomster av hendelsen A er Poisson-fordelt hvis:
- 1. Antallet av A i disjunkte tidsintervall er uavhengige
 - 2. Forventa antall av A er konstant lik λ (raten) per tidsenhet
 - 3. Kan ikke få to forekomster samtidig
 - $X =$ antall forekomster av A i et tidsrom t
 - $f(x) = \frac{(\lambda t)^x \cdot e^{-\lambda t}}{x!}, \ x = 0, 1, 2, \dots$
 - $E(X) = \lambda t \quad Var(X) = \lambda t$
 - $P(X \leq x) = \sum_{k=0}^x P(X = k)$
 - Ventetida til hendelse k er gammafordelt med $\alpha = k$ og $\beta = 1/\lambda$
 - Ventetida til første hendelse, og mellom etterfølgende hendelser, er eksponensialfordelt

Uniform fordeling:

- En kontinuerlig uniformt fordelt variabel, har samme sannsynlighet for alle verdier innen et intervall. Generelt har vi tetthetsfunksjonen: $f(x) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B \\ 0, & \text{ellers} \end{cases}$
- $E(X) = \frac{A+B}{2} \quad Var(X) = \frac{(A-B)^2}{12}$

Gammafordeling:

- En kontinuerlig variabel X er gammelfordelt med parameter $\alpha > 0$ og $\beta > 0$ dersom tetthetsfunksjonen er gitt ved: $f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x > 0 \\ 0, & \text{ellers} \end{cases}$
- $E(X) = \alpha\beta \quad Var(X) = \alpha\beta^2$

Eksponensialfordeling:

- $f(x; \beta) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}}, & x > 0 \\ 0, & \text{ellers} \end{cases}$
 - $E(X) = \beta \quad Var(X) = \beta^2$
- Eksponensialfordelinga er minneløs!

Normalfordeling:

- $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- $P(a \leq X \leq b) = \int_a^b f(x) dx$

Standard normalfordeling:

- Alle normalfordelinger kan skrives som Standard normalfordeling
- $Z = \frac{X-\mu}{\sigma}$
- $F(x) = P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Anta at X_1, X_2, \dots, X_n er uavhengige og normalfordelt. Da er: $Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n$ Være normalfordelt med:

- $E(Y) = \sum_{i=1}^n \alpha_i \mu_i \quad Var(Y) = \sum_{i=1}^n \alpha_i^2 \sigma_i^2$

Inferens:

QQ-Plot:

- Plotter observasjoner mot teoretiske(”ideelle”) observasjoner fra en aktuell fordeling.
- Teoretiske observasjoner er gitt ved invers kumulativ fordeling av jevnt spredte Sannsynlighetsfordelinger mellom 0 og 1.
- Om antatt fordeling stemmer skal plottet gi tilnærmet rett linje.

Estimering:

Viktige estimatoregenskaper:

- En punktestimator Θ for en paramaeter θ er forventningsrett hvis $E(\Theta) = \theta$
- Variansen $Var(\Theta)$ burde synke med økende antall observasjoner
- Om en har to ulike estimatore, så er den estimatoren med minst varians den mest effektive estimatoren.

Vanlige estimatorer:

Alle estimatorene vist til her er forventningsrett.

- μ : $\overline{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad E(\overline{X}) = \mu \quad Var(\overline{X}) = \frac{\sigma^2}{n}$
- σ^2 : $S^2 = \frac{1}{1-n} \sum_{i=1}^n (X_i - \overline{X})^2 \quad E(S^2) = \sigma^2 \quad Var(S^2) = \frac{2\sigma^4}{n-1}$
- p : $\hat{p} = \frac{X}{n} \quad E(\hat{p}) = p \quad Var(\hat{p}) = \frac{p(1-p)}{n}$ Binomisk
- $\mu_1 - \mu_2$: $\overline{X}_1 - \overline{X}_2 \quad Var(\overline{X}_1 - \overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
- $\frac{\sigma_1^2}{\sigma_2^2}$: $\frac{S_1^2}{S_2^2}$
- $p_1 - p_2$: $\hat{p}_1 - \hat{p}_2$, Binomisk
- μ_D : \overline{D}

Utvalgsfordelinger: