Formelsamling (Sannolikhetsteori 1)

Diskreta fördelningar

Binomialfördelning: $X \sim Bin(n,p)$

- frekvensfunktionen $P(X=x) = \binom{n}{x} p^x(1-p)^{n-x}, \quad x = 0,1,...,n$
- väntevärdet np
- variansen $np(1-p)$
- momentgenererande funktionen $M_X(t) = (pe^t + 1 - p)^n$

Poissonfördelning: $X \sim Pois(\lambda)$

- frekvensfunktionen $P(X=x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0,1,2,...$
- väntevärdet λ
- variansen λ
- momentgenererande funktionen $M_X(t) = \exp(\lambda(e^t - 1))$

Geometrisk fördelning: $X \sim Geom(p)$

- frekvensfunktionen $P(X=x) = (1-p)^{x-1}p, \quad x = 1,2,...$
- väntevärdet $\frac{1}{p}$
- variansen $\frac{1-p}{p^2}$

Negativ binomialfördelning: $X \sim NegBin(r,p)$

- frekvensfunktionen $P(X=x) = \binom{x-1}{r-1} p^r(1-p)^{x-r}, x = r, r+1, ...$
- väntevärdet $\frac{r}{p}$
- variansen $\frac{r(1-p)}{p^2}$

Hypergeometrisk fördelning: $X \sim HypGeom(n,N,m)$

- frekvensfunktionen $P(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \quad x = 0,1,...,n$
- väntevärdet $\frac{nm}{N}$
- variansen $\frac{nm}{N}\left(\frac{(n-1)(m-1)}{N-1} + 1 - \frac{nm}{N}\right)$

Kontinuerliga fördelningar

Likformig fördelning: $X \sim Lik(a, b)$

- täthetsfunktionen $f(x) = \frac{1}{b-a}, \quad a < x < b$
- väntevärdet $\frac{1}{2}(a+b)$
- variansen $\frac{1}{12}(b-a)^2$
- momentgenererande funktionen $M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$

Normalfördelning: $X \sim N(\mu, \sigma^2)$

- täthetsfunktionen $f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad -\infty < x < \infty$
- väntevärdet μ
- variansen σ^2
- momentgenererande funktionen $M_X(t) = \exp(\mu t + \sigma^2 t^2/2)$

Standardiserad normalfördelning: $Z \sim N(0, 1)$

- täthetsfunktionen $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$
- momentgenererande funktionen $M_Z(t) = \exp(t^2/2)$

Exponentialfördelning: $X \sim Exp(\lambda)$

- täthetsfunktionen $f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$
- väntevärdet $\frac{1}{\lambda}$
- variansen $\frac{1}{\lambda^2}$
- momentgenererande funktionen $M_X(t) = \frac{\lambda}{\lambda - t}$

Gammafördelning: $X \sim Gamma(s, \lambda)$

- täthetsfunktionen $f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{s-1}}{\Gamma(s)}, \quad x \geq 0$, where $\Gamma(s) = \int_0^\infty e^{-y} y^{s-1}\, dy$
- väntevärdet $\frac{s}{\lambda}$
- variansen $\frac{s}{\lambda^2}$
- momentgenererande funktionen $M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^s$

Exempel om hur olika fördelningar kan tillämpas:

- **Binomial:** Antal succéer i n oberoende försök, där succésannolikheten är p, är binomialfördelat.
- **Poisson:** Kan användas som fördelning för antal punkter i en stokastisk punktprocess (under några enkla antaganden). Approximerar binomialfördelningen, när n är stor och p liten, $\lambda = np$.
- **Geometrisk:** Antal försök, som behövs tills en händelse med sannolikhet p inträffar, har geometrisk fördelning.
- **Negativ binomial:** Antal försök, som behövs tills en händelse med sannolikhet p inträffar för r-te gången, har negativ binomialfördelning.
- **Hypergeometrisk:** Man använder fördelningen om man drar utan återläggning från en ändlig population som har två olika slags individer.
- **Likformig:** Används som fördelning för väntetider och avrundning av mätningsfel.
- **Normal:** Under generella antaganden är en summa av ett stort antal stokastiska variabler approximativt normalfördelat (centrala gränsvärdessatsen).
- **Standardiserad normal:** Om $X \sim N(\mu, \sigma)$, då har $\frac{X-\mu}{\sigma}$ standardiserad normalfördelning.
- **Exponential:** Fördelning för livslängd (utan åldrande).
- **Gamma:** Fördelning för summan av n oberoende stokastiska variabler, som är exponentialfördelade med parameter λ.

Summor av oberoende stokastiska variabler:

- $X_1 \sim Bin(n_1, p)$ och $X_2 \sim Bin(n_2, p)$: $X_1 + X_2 \sim Bin(n_1 + n_2, p)$
- $X_1 \sim Pois(\lambda_1)$ och $X_2 \sim Pois(\lambda_2)$: $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$
- $X_1 \sim Exp(\lambda)$ och $X_2 \sim Exp(\lambda)$: $X_1 + X_2 \sim Gamma(2, \lambda)$
- $X_1 \sim Gamma(n_1, \lambda)$ och $X_2 \sim Gamma(n_2, \lambda)$:
  $X_1 + X_2 \sim Gamma(n_1 + n_2, \lambda)$
- $X_1 \sim N(\mu_1, \sigma_1^2)$ och $X_2 \sim N(\mu_2, \sigma_2^2)$: $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Bivariat normalfördelning: En två-dimensionell stokastisk variabel (X, Y) har en bivariat normalfördelning med parametrar μ_1, μ_2, σ_1^2, σ_2^2 och ρ, om dess täthetsfunktion är

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot$$
$$\exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right)\right).$$

Den betingade fördelningen av Y givet $X = x$ är normalfördelning med väntevärde $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$ och varians $\sigma_2^2(1-\rho^2)$. Om (X, Y) har en bivariat normalfördelning med parametrar μ_1, μ_2, σ_1^2, σ_2^2 och ρ, då har $\left(\frac{X-\mu_1}{\sigma_1}, \frac{Y-\mu_2}{\sigma_2}\right)$ bivariat normalfördelning med parametrar 0, 0, 1, 1 och ρ.

Svaga versionen av stora talens lag: Låt $(X_k : k \geq 1)$ vara oberoende och lika fördelade stokastiska variabler med ändligt väntevärde μ. Då är för varje $\epsilon > 0$, då $n \to \infty$

$$P\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu\right| > \epsilon\right) \to 0.$$

Centrala gränsvärdessatsen: Låt $(X_k : k \geq 1)$ vara oberoende och lika fördelade stokastiska variabler med väntevärde μ och ändlig varians σ^2. Då är, då $n \to \infty$

$$P\left(\frac{X_1 + X_2 + ... + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right) \to \Phi(a),$$

där Φ är fördelningsfunktionen av den standardiserade normalfördelningen.

Obs! Normalfördelningstabellen delas ut också.