Föreläsning 7 sammanfattning

Litteratur

Delar från kursboken 6.2-6.3, 6.5, 6.7

Normalfördelning

Detta är den viktigaste av alla fördelningar. Kallas ibland för klockkurvan.

Definition

En kontinuerlig s.v. X sägs vara normalfördelad med parametrar μ och $\sigma, (\sigma > 0)$ om täthetsfunktionen ges av

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Beteckning: $X \in N(\mu, \sigma)$.

Anmärkning:

- Effekten av att ändra μ är att täthetens läge förskjuts.
- Effekten av att ändra σ är att fördelningen blir mer eller mindre koncentrerad kring masscentrat.

Normalfördelningens fördelningsfunktion

$F_X(x)$ ges av

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Integralen har inget slutet uttryck. För givna μ, σ och x kan den beräknas numeriskt.

Standardiserad normalfördelning

Om en normalfördelad s.v. Z har parametrarna $\mu = 0$ och $\sigma = 1$ sägs den vara *standardiserat normalfördelad,* $Z \in N(0,1)$. Detta specialfall av normalfördelningen har egna beteckningar för fördelningsfunktionen $F_Z(z) = \Phi(z)$ och täthetsfunktionen $f_Z(z) = \phi_Z(z)$. Dessa definieras enligt

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty,$$

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{t}} dt.$$

Senare kommer vi se att det räcker med att kunna räkna ut dessa funktioner för att kunna bärkna $F_X(x)$ för en godtycklig normalfördelning.

Egenskaper för standardiserad normalfördelning

- $\phi(-z) = \phi(z), \quad \Phi(-z) = 1 - \Phi(z)$
- För $Z \in N(0,1)$ gäller att

  $$P(a < Z \leq b) = \Phi(b) - \Phi(a)$$

  Och eftersom att fördelningen är kontinuerlig kan $<$ bytas mot $\leq$ godtyckligt eller omvänt utan att sannolikheten ändras vilket kan vara bra att ha i åtanke om man stöter på ett uttryck där $<$ förekommer men där det är mer praktiskt att räkna med $\leq$ exempelvis.
- Kvantiler för $N(0,1)$ förekommer så ofta att dessa givits en egen beteckning: λ_α.
- α-kvantilen för en standardiserad normalfördelning definieras som löningen till $P(Z > \lambda_\alpha) = \alpha$. Men då $P(Z > \lambda_\alpha) = 1 - \Phi(\lambda_\alpha)$ så löser tydligen λ_α

  $$\Phi(\lambda_\alpha) = 1 - \alpha$$
- Om $Z \in N(0,1)$ så gäller att

  $$E(Z) = 0, \quad D(Z) = 1.$$

  Bevis: Eftersom $\phi(\cdot)$ är symmetrisk kring 0 så får man

  $$E(Z) = \int_{-\infty}^{\infty} z\phi(z)dz = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0$$

- För att få $D(Z)$ använder vi satsen som säger att $V(Z) = E(Z^2) - (E(Z))^2$ och genom partiell integration fås

$$E(Z^2) = \int_{-\infty}^{\infty} z^2 \phi(z) dz = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$= \left[-z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0 + 1 = 1.$$

- $X \in N(\mu, \sigma)$ omm $Z = \frac{X-\mu}{\sigma} \in N(0,1)$.

- Tolkning av μ och σ:

$$E(X) = E(\sigma Z + \mu) = \sigma E(Z) + \mu = \mu,$$

$$V(X) = V(\sigma Z + \mu) = \sigma^2 V(Z) = \sigma^2$$

  D.v.s. parametrarna μ och σ är väntevärde respektive standardavvikelse för en $N(\mu, \sigma)$-fördelad s.v.

- Vidare gäller att

$$f_X(x) = \frac{1}{\sigma} \phi(\frac{x-\mu}{\sigma}), F_X(x) = \Phi(\frac{x-\mu}{\sigma}),$$

$$P(a < X \leq b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma}).$$

  Detta resultat kan vara mycket användbart.

Centrala gränsvärdessatsen (CGS)

Den viktigaste satsen inom sannolikhetsteorin och säger att en summa av oberoende lika fördelade s.v. med godtycklig fördelning är ungefär normalfördelan så länge antalet komponenter i summan är tillräckligt stort.

Definition

$X_1, ..., X_n, ...$ är en oändlig följd oberoende likafördelade s.v. med väntevärde μ och standardavvikelse $0 < \sigma < \infty$. Därav

$$Y_n = X_1 0 + ... + X_n.$$

Då gäller för givna $a < b$ att

$$P\left(a < \frac{Y_n - n\mu}{\sqrt{n}\sigma} \leq b \right) \to \Phi(b) - \Phi(a), \quad \text{då } n \to \infty$$

CGS uttalar sig alltså om fördelningen av Y_n då antalet n växer mot oändligheten: Y_n är ungefär $N(n\mu, \sqrt{n}\sigma)$-fördelad.

Beteckning: $Y_n \in AsN(n\mu, \sqrt{n}\sigma)$

- Observera att $E(Y_n) = n\mu$ och $D(Y_n) = \sqrt{n}\sigma$. För varje givet n är

  $$\frac{Y_n - n\mu}{\sqrt{n}\sigma}$$

  en standardiserad s.v. Den har väntevärde lika med noll och en standardavvikelse lika med 1 som en standardiserad normalfördelad s.v.

- Enligt CGS: när n går mot oändligheten kommer hela fördelningen för den angivna standardiserade s.v. att gå mot en standardiserad normalfördelning, d.v.s.

  $$\frac{Y_n - n\mu}{\sqrt{n}\sigma} \in AsN(0,1).$$

- **Följdsats:** För en oändlig följd av oberoende likafördelade s.v. $X_1, ..., X_n, ...$ med $E(X_i) = \mu$ och $D(X_i) = \sigma (0 < \sigma < \infty)$ gäller att

  $$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \in AsN(\mu, \frac{\sigma}{\sqrt{n}}), \quad \text{då } n \to \infty.$$

  D.v.s det arigmetiska medelvärdet $\bar{X}$ är approximativt normalfördelat för tillräckligt stort n.

- Normalfördelningsapproximation. Enligt CGS: $\sum_{i=1}^{n} X_i \in AxN(n\mu, \sqrt{n}\sigma)$ och $\bar{X} \in AsN(\mu, \sigma/\sqrt{n})$. Detta ger approximationerna

  $$P\Big(a < \sum_{i=1}^{n} X_i \leq b\Big) \approx \Phi\Big(\frac{b - n\mu}{\sqrt{n}\sigma}\Big) - \Phi\Big(\frac{a - n\mu}{\sqrt{n}\sigma}\Big),$$

  $$P(c < \bar{X} \leq d) \approx \Phi\Big(\frac{d - \mu}{\sigma/\sqrt{n}}\Big) - \Phi\Big(\frac{c - \mu}{\sigma/\sqrt{n}}\Big).$$