Föreläsning 4 sammanfattning

Litteratur

Från kursboken 3.6-3.7, 3.10 & 5.3

Kontinuerliga stokastiska variabler

En kontinuerlig s.v. antar *alla* värden i ett eller flera intervall av R^1 . Det finns med andra ord ett oändligt antal värden som X kan anta. Detta är i praktiken ofta inte tillämpbart då man använder sig av mätinstrument med begränsad upplösning men det är ändå praktiskt att tänka sig att det är på detta vis.

Därmed så ligger utfallen utspridda längs x-axeln oändligt tätt och inget enskilt utfall kan antas med positiv sannolikhet. Vi behöver därmed definiera sannolikhetsfunktionen annorlunda än tidigare.

Vi tänker oss att sannolikhetsmassan 1 ligger utspridd på R^1 enligt *täthetsfunktionen* $f_X(x), x \in R^1$.

För händelsen då $X \in A$ fås sannolikheten för händelsen enligt

$$P(X \in A) = \int_A f_X(x)dx.$$

Villkoren för att $f_X(x)$ ska vara en täthetsfunktion för X är

- $f_X(x) \geq 0$,
- $\int_{-\infty}^{\infty} f_X(x)dx = 1$, dvs hela arean under täthetsfunktionen är 1.

Vanliga fördelningar för kontinuerliga s.v.

Likformig fördelning

Gäller om en s.v. X har täthetsfunktionen

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{om} \quad a < x < b \\ 0 & \text{f.ö.} \end{cases}$$

Betecknas enl. $X \in U(a, b)$ där U kommer från engelskans *uniform*.

Det krävs att man använder sig av intervall och inte enskilda punkter då $P(X = x) = 0$.

Exponentialfördelning

Om en s.v. X har täthetsfunktion

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & om \quad x > 0, \\ 0 & \text{f.ö.} \end{cases}$$

Används ibalnd för att modellera livslängden på elektroniska komponenter. Fördelningen har egenskapen att den saknar minne. Så en elektrisk komponent med exponensialfördelad livslängd som har hållt i t tidsenheter har samma sannolikhet att hålla lika lång tid till. Andra tillämpningar är tiden tills någon får sitt nästa telefonsamtal, tiden tills någon råkar ut för sin nästa bilolycka eller avståndet mellan mutationer på en DNA-sträng.

Normalfördelning

En s.v. X som har täthetsfunktionen

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

där μ och $\sigma > 0$ är givna tal, sägs vara normalfördelad. Denna fördelning har beteckning $X \in N(\mu, \sigma)$.

Denna fördelning är den viktigaste. Kallas även för Gauss-fördelning. Att den är så viktig beror på att om man summerar många stokastiska variabler (tänk: något slumpmässigt som beror på många olika faktorer) så är resultatet nästan alltid normalfördelat. Denna fördelning återkommer i frl 7 och beskrivs i boken i kap. 6.

Väntevärde och varians

Väntevärdet för en kontinuerlig s.v. X definieras enl.

$$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

För en ytterligare s.v. Y som är en funktion av den s.v. X enligt $Y = g(X)$ där $g(\cdot)$ är en reellvärd funktion definieras Väntevärdet enligt

$$E(Y) = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

Variansen av en kontinuerlig s.v. X definieras enl.

$$\sigma_X = V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx.$$

Ytterligare så gäller även att

$$V(X) = E(X^2) - (E(X))^2$$

precis som i det diskreta fallet.

Fördelningsfunktion för en kontinuerlig s.v.

Funktionen

$$F_X(x) = P(X \leq x), x \in R^1$$

kallas för fördelningsfunktionen för den kontinuerliga s.v. X.

Villkor

- $0 \leq F_X(x) \leq 1$
- $F_X(x)$ är icke-avtagande funktion av x,
- $F_X(x) \to 0$ då $x \to -\infty$ och $F_X(x) \to 1$ då $x \to \infty$
- $F_X(x)$ är kontinuerlig till höger för varje x .

Det förekommer ofta att man behöver beräkna sannolikheter på formen $P(a < X \leq b), P(X \leq a), P(X > a)$, eller liknande. I dessa fall så är fördelningsfunktionen $F_X(x)$ mycket användbar.

Som ett exempel så gäller att om $a < b$ så

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

Ytterligare så gäller vid beräkning av s.k. kvantiler att om α-kvantilen söks. Så fås denna av lösningen $x = x_\alpha$ till ekvationen

$$F_X(x) = 1 - \alpha$$