4. Kontinuerliga fördelningar

En stokastisk variabel är **kontinuerlig** om den kan anta vilket värde som helst på ett (eller flera) intervall i $\mathbf{R}$.

För att räkna sannolikheter använder man **täthetsfunktionen** (density function) av den stokastiska variabeln X. För en täthetsfunktion gäller det att

1) $f(x) \geq 0$ för $x \in \mathbf{R}$

2) $\int_{-\infty}^{\infty} f(x)\, dx = 1$

3) $P(a \leq X \leq b) = \int_a^b f(x)\, dx$ för $a, b \in \mathbf{R}$, $a \leq b$

En funktion f är en täthetsfunktion om och endast om 1) och 2) gäller.

Obs! Från 3) får man att i kontinuerliga fallet är sannolikheten av vilket som helst specifikt värde lika med noll. T. ex. $P(X = a) = P(a \leq X \leq a) = \int_a^a f(x)\, dx = 0$.

Ex: Blykonsentrationen i bensin varierar mellan 0.1g/l och 0.5g/l. Man vill räkna sannolikheten att blykonsentrationen i en bensinliter (tagen på måfå) är mellan 0.2g och 0.3g. För att kunna räkna sannolikheten antar man att X är mängden av bly (i gram) i en bensinliter och att

$$f(x) = \begin{cases} 12.5x - 1.25, & 0.1 \leq x \leq 0.5 \\ 0, & \text{annars} \end{cases}$$

Obs!

1) Om man bara skriver

$$f(x) = 12.5x - 1.25, \quad 0.1 \leq x \leq 0.5$$

betyder det att annars är $f(x) = 0$.

2) Den totala arean under täthetsfunktionskurvan är lika med 1. Sannolikheten att X är mellan a och b är arean av den delen av kurvan som är mellan a och b.

3) $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$ därför att $P(X = a) = P(X = b) = 0$.

För en kontinuerlig stokastisk variabel X är **fördelningsfunktionen**

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)\, dy, \quad x \in \mathbf{R}$$

I kontinuerliga fallet är $f(x) = F'(x)$.

4.2 Väntevärdet och fördelningsparametrar

Väntevärdet av en kontinuerlig stokastisk variabel X är

$$\mathbf{E}[X] = \mu = \int_{-\infty}^{\infty} x f(x)\, dx.$$

Om $H(X)$ är en funktion av X, är

$$\mathbf{E}[H(X)] = \int_{-\infty}^{\infty} H(x) f(x)\, dx.$$

Variansen är (som i diskreta fallet)

$$Var(X) = \sigma^2 = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - \mu^2$$

4.4 Normalfördelning

En stokastisk variabel med täthetsfunktionen

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty,$$

där $\mu \in \mathbf{R}$ och $\sigma > 0$, är normalfördelad med parametrar μ och σ, $X \sim N(\mu, \sigma)$.

Parametern μ är **väntevärdet** av X och σ är **standardavvikelsen** av X.

Ex. 4.4.1: En av de viktigaste orsaken för föroreningen av luften är kolväte (hydrocarbon) från bilar. Låt den stokastiska variabeln X vara mängden av kolväte (i gram) per km från en bil. Antag att X är normalfördelad med väntevärdet 1g och standardavvikelsen 0.25g.

Sats 4.4.3: Låt X vara normalfördelad med väntevärdet μ och standardavvikelsen σ. Då har den stokastiska variabeln $\frac{X-\mu}{\sigma}$ så-kallad standardiserad normalfördelning, dvs. normalfördelning med väntevärdet $\mu = 0$ och standardavvikelsen $\sigma = 1$.

Ex. 4.4.2: (Ex 4.4.1 fortsätter) Vad är sannolikheten att en bil vald på måfå släpper mellan 0.9g och 1.54g kolväte per km?

Ex 4.4.3: Låt X vara mängden av bestrålning man kan absorbera innan man dör av bestrålningen. Anta att $X \sim N(500, 150)$. Över vilken bestrålningsnivå kan bara 5% av dem som blir bestrålade överleva?