5 Kontinuerliga stokastiska variabler

Ex: X är livslängden av en glödlampa. Utfallsrummet är $S = \{x : x \geq 0\}$. X kan anta överuppräkneligt oändligt många olika värden. X är en kontinuerlig stokastisk variabel.

Definition: X är en kontinuerlig stokastisk variabel om det finns en ickenegativ funktion f, som är definierad för alla $x \in (-\infty, \infty)$, sådan att för vilken som helst mängd $B \subset \mathbf{R}$

$$P(X \in B) = \int_B f(x)\, dx.$$

Funktionen f kallas *täthetsfunktionen* (density function) av X.

Ex.1a: X är en kontinuerlig stokastisk variabel med täthetsfunktionen

$$f(x) = \begin{cases} c(4x - 2x^2), & 0 < x < 2 \\ 0 & \text{annars} \end{cases}.$$

a) Vad måste värdet av c vara?

b) Beräkna $P(X > 1)$.

Ex.1c: Livslängden X av en radiorör är en stokastisk variabel med täthetsfunktionen

$$f(x) = \begin{cases} 0, & x \leq 100 \\ \frac{100}{x^2}, & x > 100 \end{cases}$$

Hur stor är sannolikheten att exakt två av fem sådana rör måste bytas under de första 150 timmarna, som radion är på? Anta att händelserna E_i = det i-te röret måste bytas, $i = 1, 2, ..., 5$, är oberoende.

Fördelningsfunktionen av den stokastiska variabeln X är

$$F(a) = P(X \leq a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)\, dx.$$

Täthetsfunktionen är derivatan av fördelningsfunktionen, dvs. $F'(a) = f(a)$.

5.2 Väntevärdet och variansen av en kontinuerlig stokastisk variabel

I kontinuerliga fallet är väntevärdet av X

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

Ex.2a: Vad är väntevärdet av den stokastiska variabeln X, som har täthetsfunktionen

$$f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0 & \text{annars} \end{cases} \quad ?$$

Ex: Låt X vara en stokastisk variabel med täthetsfunktionen

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0 & \text{annars} \end{cases}$$

Vad är väntevärdet av $\mathbf{E}[X^2]$?

Proposition 2.1: Om X är en kontinuerlig variabel med täthetsfunktionen $f(x)$, då för vilken som helst funktion g ($g(x) \in \mathbf{R}$)

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

För att visa propositionen behöver man den följande lemman

Lemma 2.1: För en icke-negativ stokastisk variabel Y

$$\mathbf{E}[Y] = \int_{-\infty}^{\infty} P(Y > y)\, dy.$$

Ex.2c: En pinne av längd 1m är bruten i två bitar. Punkten U, där pinnen är bruten, har likformig fördelning på intervallet $(0, 1)$, dvs. att $f_U(u) = 1$, om $x \in (0, 1)$, och $f_U(u) = 0$ annars. Vad är den förväntade längden av den delen av pinnen som innehåller en viss punkt p?

Följdsats 2.1: Om a och b är konstanter, då är

$$\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$$

Variansen av en kontinuerlig stokastisk variabel definieras på samma sätt som i diskreta fallet, dvs. att

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

Obs! $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$ (som i diskreta fallet).

5.3 Likformig fördelning

En stokastisk variabel har likformig fördelning på intervallet (α, β) om den har täthetsfunktionen

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & \alpha < x < \beta \\ 0 & \text{annars} \end{cases}$$

Man skriver $X \sim Lik(\alpha, \beta)$. Väntevärdet av X är $\frac{1}{2}(\alpha + \beta)$ och variansen är $\frac{1}{12}(\beta - \alpha)^2$.

Fördelningsfunktionen av $Lik(\alpha, \beta)$ är

$$F(x) = \begin{cases} 0, & x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha}, & \alpha < x < \beta \\ 1 & x \geq \beta \end{cases}$$

Ex. 3c: Bussar kommer till en hållplats varje 15 minuter f.o.m. kl 7:00. Om en passagerare kommer till hållplatsen vid en tid som är likformigt fördelad mellan 7:00 och 7:30, vad är sannolikheten att hon väntar

a) mindre än 5 minuter?

b) längre än 10 minuter?

5.4 Normalfördelning

Definition: Om den stokastiska variabeln X har täthetsfunktionen

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

där μ och σ är givna storheter ($\sigma > 0$), har X en *normalfördelning* med parametrar μ (väntevärde) och σ^2 (varians), $X \sim N(\mu, \sigma^2)$.

Om $X \sim N(\mu, \sigma^2)$, då är

$$\alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2).$$

Speciellt är

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

Man säger att Z har den *standardiserade normalfördelningen.*

Ex.4d: På en rättegång vittnar en expert att längden av graviditeten är approximativt normalfördelad med parametrar $\mu = 270$ och $\sigma^2 = 100$. En man (som bekyller sin fru) kan bevisa att han var utomlands under en period som började 290 dagar och slutade 240 dagar innan förlossningen. Om mannen är fadern (vilket han inte tror på), vad är sannolikheten att moderns graviditet skulle ha varit väldigt lång (längre än 290 dagar) eller väldigt kort (mindre än 240 dagar)?

5.5 Exponentialfördelningen

En kontinuerlig stokastisk variabel dess täthetsfunktion är

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & \text{annars} \end{cases}$$

för någon $\lambda > 0$ är en exponentialfördelad med parameter λ, $X \sim Exp(\lambda)$. Väntevärdet av X är $\frac{1}{\lambda}$ och variansen $\frac{1}{\lambda^2}$.

Ex.5b: Anta att längden av ett telefonsamtal (i minuter) är exponentialfördelad med parameter $\lambda = \frac{1}{10}$. Om någon kommer just innan dig till en telefonkiosk, vad är sannolikheten att du måste vänta

a) mer än 10 minuter?

b) mellan 10 och 20 minuter?

Glömske-egenskapen av exponentialfördelningen: Låt $X \sim Exp(\lambda)$. Då är

$$P(X > s + t | X > t) = P(X > s) \quad \text{för varje } s, t \geq 0.$$

Ex.5d: Anta att antalet kilometer man kan köra en bil innan batteriet är slut är exponentialfördelat med genomsnitt 10,000km. Om man önskar att köra 5,000km, hur stor är sannolikheten att man inte behöver byta batteriet under resan? Vad kan man säga om sannolikheten om man inte antar exponentialfördelningen?

5.6 Andra kontinuerliga fördelningar

5.6.1 Gammafördelning

En stokastisk variabel har en *gammafördelning* med parametrar (α, λ), $\lambda > 0$, $\alpha > 0$, om dess täthetsfunktion är

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & x \geq 0 \\ 0 & x < 0 \end{cases},$$

där gammafunktionen $\Gamma(\alpha)$ definieras

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy.$$

Man skriver $X \sim Gamma(\alpha, \lambda)$. Väntevärdet av X är $\frac{\alpha}{\lambda}$ och variansen $\frac{\alpha}{\lambda^2}$.

5.7 Fördelning av en funktion av en stokastisk variabel

Man känner till fördelningen för X och skulle vilja härleda fördelningen för någon funktion av X, $g(X)$.

Ex.7a: Låt $X \sim Lik(0,2)$ och $Y = X^2$. Vad är fördelningen för Y?

Sats 7.1: Låt X vara en kontinuerlig stokastisk variabel med täthetsfunktionen f_X. Anta att g är en strängt monoton (växande eller avtagande), deriverbar (och kontinuerlig) funktion av X. Då är täthetsfunktionen av $Y = g(X)$

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))|\frac{d}{dy}g^{-1}(y)|, & \text{om } y = g(x) \text{ för någon } x \\ 0 & \text{om } y \neq g(x) \text{ för varje } x \end{cases},$$

där $g^{-1}(y)$ är definierad så att den är lika med värdet av x sådant att $g(x) = y$.

Ex.7d: Låt X vara en ickenegativ stokastisk variabel med täthetsfunktionen f och $Y = X^n$. Vad är f_Y?