Kontinuerliga stokastiska variabler

Kom ihåg att om värdemängden Ω för en stokastisk variabel ξ är ett kontinuerligt utfallsrum, t.ex. $\Omega = [0,1]$ eller $\Omega = \mathbb{R}$, så kallas ξ för en <u>kontinuerlig stokastisk variabel.</u>

Notera att ξ kan anta oändligt många värden här.

<u>Def</u> Om fördelningen $P(A) = \mathbb{P}(\xi \in A)$, $A \subseteq \Omega$, uppfyller $\mathbb{P}(\xi \in A) = \int_A f(t)\,dt$, och i synnerhet fördelningsfunktionen ges av $F(x) = \mathbb{P}(\xi \leq x) = \int_{-\infty}^{x} f(t)\,dt$, så säger vi att ξ har <u>täthetsfunktionen/frekvensfunktionen</u> $f(t)$, $t \in \mathbb{R}$

Det gäller att:

Villkor:
- $f(x) \geq 0$ för alla $x \in \mathbb{R}$ (i synnerhet $f(x) = 0$ för $x \in \mathbb{R} \setminus \Omega = \Omega^c$)
- $P(\mathbb{R}) = \mathbb{P}(\xi \in \mathbb{R}) = \int_{-\infty}^{\infty} f(t)\,dt = 1$

- $\mathbb{P}(a < \xi < b) = \mathbb{P}(a \leq \xi < b) = \mathbb{P}(a < \xi \leq b) = \mathbb{P}(a \leq \xi \leq b)$ $= \int_a^b f(t)\,dt$, $a < b$, eftersom $\mathbb{P}(\xi = x) = \lim_{a \to x} \int_a^x f(t)\,dt = 0$ för alla $x \in \mathbb{R}$ (sannolikheten att ξ antar ett specifikt tal, t.ex. 3.987209..., är 0)

Notera även att $\mathbb{P}(\xi > x) = 1 - \mathbb{P}(\xi \leq x) =$
$= 1 - F(x) = 1 - \int_{-\infty}^{x} f(t)\,dt$

samt att $f(x)$ är derivatan av $F(x)$
(enligt analysens huvudsats).

Ex. Visa att

$$f(x) = \begin{cases} \frac{x^3}{4} & \text{om } 0 < x \leq 2 \\ 0 & \text{i övrigt} \end{cases}$$

är en täthetsfunktion, bestäm den tillhörande fördelningsfunktionen och beräkna $\mathbb{P}(\xi < 1)$ för den tillhörande stokastiska variabeln.

Lösn:
Eftersom $\frac{x^3}{4} > 0$ för alla $x \in (0,2]$ så gäller att $f(x) \geq 0$ för alla x. Vidare gäller

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^2 \frac{x^3}{4}\,dx + \int_{-\infty}^{0} 0\,dx + \int_2^{\infty} 0\,dx = \left[\frac{1}{4}\,\frac{x^4}{4}\right]_0^2 = \frac{2^4}{16} - 0 = 1$$

så $f(x)$ uppfyller de två villkoren för att vara en täthetsfunktion.

För $x \leq 0$ får vi $F(x) = \int_{-\infty}^{x} 0\,dx = 0$.

För $x \in (0,2]$ får vi $F(x) = \int_{-\infty}^{0} 0\,dx + \int_0^x \frac{t^3}{4}\,dt = \left[\frac{t^4}{16}\right]_0^x = \frac{x^4}{16}$.

För $x > 2$ får vi $F(x) = \int_2^x 0\,dx = 0$.

$\mathbb{P}(\xi < 1) = \mathbb{P}(\xi \leq 1) = F(1) = \frac{1}{16}$

Ex Antag att livslängden hos ett tändstift, ξ, i timmar, är en stokastisk variabel med tätheten

$$f(x) = \begin{cases} c e^{-x/120}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

a) Finn sannolikheten att ett tändstift fungerar mellan 100 och 150 timmar innan det går sönder.

b) Vad är sannolikheten att det fungerar längre än 200 timmar?

Lösn:

C är okänd och måste beräknas:

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = C\int_{-\infty}^{\infty} e^{-x/150}\,dx = C\left[-120 e^{-x/120}\right]_0^{\infty} = 120C$$

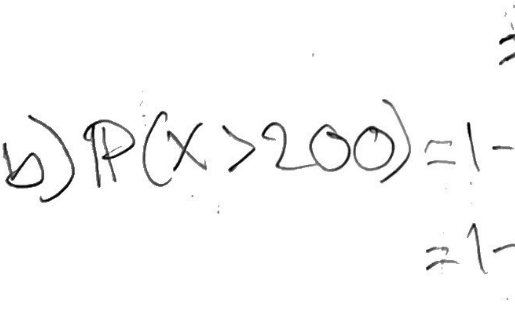
$$\Rightarrow C = \frac{1}{120}$$

a) $$P(100 < \xi < 150) = \int_{100}^{150} \frac{1}{120} e^{-x/120}\,dx = \left[-e^{-x/120}\right]_0^{150} = -e^{-5/4} + e^{-5/6}$$

$$\approx 0{,}148$$

b) $$P(X > 200) = 1 - F(200) = 1 - \int_0^{200} \frac{1}{120} e^{-x/120}\,dx = 1 - \left[-e^{-x/120}\right]_0^{200}$$

$$= 1 - (-e^{-5/3} - (-1)) = e^{-5/3} \approx 0.19$$



<u>Ex</u> Antag att livslängden i timmar, ξ, för en fiberoptisk kabel beskrivs av en stokastisk variabel med täthet

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

Finn sannolikheten att exakt tre av fem sådana kablar måste ersättas inom 150 timmars användande, om vi antar att kablarna går sönder oberoende av varandra.

<u>Lösn:</u>

Låt Y = antal kablar som måste ersättas inom 150 timmars användande

$\Rightarrow Y \sim \mathrm{Bin}(n,p)$ med $n=150$ och p given

av $p = \mathbb{P}(\xi < 150) = \int_{100}^{150} \frac{100}{x^2}\,dx = \left[-\frac{100}{x}\right]_{100}^{150} = -\frac{2}{3} - (-1) = \frac{1}{3}$

$$\Rightarrow \mathbb{P}(Y=3) = \binom{5}{3}\left(\tfrac{1}{3}\right)^3\left(1-\tfrac{1}{3}\right)^{5-3} = \frac{40}{243} \approx 0.165$$

$\boxtimes$

<u>Def</u>

För en kontinuerlig stokastisk variabel ξ med frekvens-/täthetsfunktion $f(x)$ ges <u>väntevärdet</u> av $\mu = \mathbb{E}[\xi] = \int_{-\infty}^{\infty} x f(x)\,dx$

Eftersom $\mathbb{E}[g(\xi)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx$ för en godtycklig funktion $g(x)$ får vi $\mathrm{Var}(\xi) = \mathbb{E}[\xi^2] - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{\infty} x f(x)\,dx\right)$

Ex Tätheten $f(x)=\begin{cases} \frac{x^3}{4} & , 0<x<2 \\ 0 & , \text{i övrigt} \end{cases}$ ger

$$\mathbb{E}[\xi] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^2 \frac{x^4}{4}\,dx = \left[\frac{1}{5}\cdot\frac{x^5}{4}\right]_0^2 = \frac{2^5}{20} = \frac{32}{20} = 1.6$$

När frekvensfunktionen är komplicerad har man ibland nytta av partialintegration:

$$\int_a^b g(x)h(x)\,dx = \left[\underset{\substack{\uparrow \\ \text{integralen} \\ \text{av } g}}{G(x)}h(x)\right]_a^b - \int_a^b G(x)\underset{\substack{\uparrow \\ \text{derivatan} \\ \text{av } h}}{h'(x)}\,dx$$

Ex Livslängden i timmar hos en sorts glödlampa, ξ, har tätheten

$$f(x) = \begin{cases} \frac{1}{100}e^{-x/100} & , x \geq 0 \\ 0 & , x<0 \end{cases}$$

Låt $g(x)=f(x)$ och $h(x)=x \Rightarrow G(x)=-e^{-x/100}$ och $h'(x)=1$

$$\mathbb{E}[\xi] = \int_0^{\infty} x f(x)\,dx = \int_0^{\infty} h(x)g(x)\,dx = \left[G(x)h(x)\right]_0^{\infty} - \int_0^{\infty} G(x)h'(x)\,dx$$

$$= \left[-e^{-x/100}\cdot x\right]_0^{\infty} + \int_0^{\infty} e^{-x/100}\,dx = \lim_{x\to\infty} \frac{-x}{e^{x/100}} + \left[-100e^{-x/100}\right]_0^{\infty}$$

$$= 0 + 0 - (-100) = 100$$

Ex Frekvensfunktionen

$$f(x) = \begin{cases} \frac{x^3}{4} & , 0<x<2 \\ 0 & , \text{i övrigt} \end{cases}$$

Vi vet att $\mathbb{E}[\xi] = \frac{8}{5}$ sedan tidigare

$$\mathbb{E}[\xi^2] = \int_0^2 x^2 f(x)\,dx = \int_0^2 x^2 \frac{x^3}{4}\,dx = \left[\frac{1}{6}\frac{x^6}{4}\right]_0^2 = \frac{64}{24}$$

$$\Rightarrow \operatorname{Var}(\xi) = \mathbb{E}[\xi^2] - \mathbb{E}[\xi]^2 = \frac{64}{24} - \left(\frac{8}{5}\right)^2 = \frac{8}{75}$$

$$\Rightarrow \sigma = \sqrt{\operatorname{Var}(\xi)} = \sqrt{\frac{8}{75}} \approx 0.3266$$

Ex Täthet för spårvagnsväntetid:

$$f(x) = \begin{cases} Cx & 0<x\leq 10 \\ 0.2-0.01x & 10<x\leq 20 \\ 0 & \text{övrigt} \end{cases}$$

a) Bestäm C

b) Bestäm fördelningsfunktionen

c) Genomsnittlig väntetid?

d) Standardavvikelsen?

e) Sannolikheten att väntetiden är längre än fem minuter?

Lösn

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = C\int_0^{10} x\,dx + \int_{10}^{20} 0.2-0.01x\,dx = 50C + (4-2) - (2-\tfrac{1}{2})$$

$$\Leftrightarrow 50C = \frac{1}{2} \Leftrightarrow C = \frac{1}{100}$$

b) $\underline{x \le 0}$: $F(x) = 0$

$\underline{0 < x \le 10}$: $F(x) = \int_0^x 0.01t\,dt = \left[\frac{0.01}{2} t^2\right]_0^x = 0.005x^2$

$\underline{10 < x < 20}$: $F(x) = 0.005 \cdot 10^2 + \int_{10}^{x} (0.2 - 0.01t)\,dt =$

$= 0.5 + \left[0.2t - \frac{0.01t^2}{2}\right]_{10}^{x} = -\frac{0.01x^2}{2} + 0.2x - 1$

$\underline{x \ge 20}$: $F(x) = 1$

c) $\mathbb{E}[\xi] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{10} x(0.01x)\,dx + \int_{10}^{20} x(0.2 - 0.01x)\,dx =$

$= \left[\frac{0.01x^3}{3}\right]_0^{10} + \left[\frac{0.2x^2}{2} - \frac{0.01x^3}{3}\right]_{10}^{20} = 10$

d) $\mathbb{E}[\xi^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^{10} x^2(0.01x)\,dx + \int_{10}^{20} x^2(0.2 - 0.01x)\,dx$

$= \left[\frac{0.01x^4}{4}\right]_0^{10} + \left[\frac{0.2x^3}{3} - \frac{0.01x^4}{4}\right]_{10}^{20} = 116\frac{2}{3}$

$\Rightarrow Var(\xi) = \mathbb{E}[\xi^2] - \mu^2 = 116\frac{2}{3} - 100 = 16\frac{2}{3}$

$\Rightarrow \sigma = S(\xi) = \sqrt{16\frac{2}{3}} \approx 4.082$

e) $P(\xi > 5) = 1 - F(5) = 1 - 0.005 \cdot 5^2 = 1 - 0.125 = 0.875$

Vanliga fördelningar

Rektangelfördelning/kontinuerlig likformig

Antag att tätheten är konstant på ett intervall $[a,b]$ och 0 utanför $[a,b]$, dvs inom $[a,b]$ är alla utfall lika (o)sannolika.

- $$1=\int_{-\infty}^{\infty} f(x)\,dx = \int_a^b c\,dx \Rightarrow f(x)=\frac{1}{b-a} \text{ för } x\in[a,b]$$

- **Def**

  ξ har rektangelfördelning/likformig kontinuerlig fördelning på $[a,b]\subseteq \mathbb{R}$

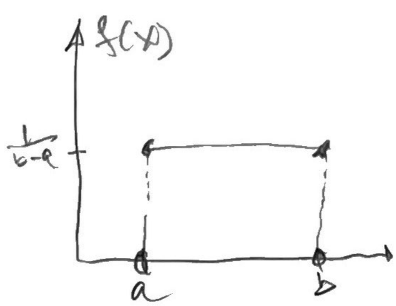
  om $$f(x)=\begin{cases}\frac{1}{b-a} & ,\ a\le x\le b\\ 0 & \text{annars}\end{cases}$$

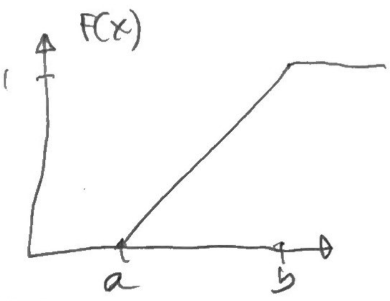
  ($\xi\sim R[a,b]$, $\xi\in R[a,b]$
  $\xi\sim Uni(a,b)$, $\xi\in Uni(a,b)$)

- $$F(x)=\int_{-\infty}^{x}\frac{1}{b-a}\,dt=\begin{cases}0 & ,\ x<a\\ \frac{x}{b-a} & ,\ a\le x\le b\\ 1 & ,\ x>b\end{cases}$$

- $$\mathbb{E}[\xi]=\int_a^b \frac{x}{b-a}\,dx=\left[\frac{x^2}{2}\right]_a^b\Big/(b-a)=\frac{b^2-a^2}{2(b-a)}=\frac{(b-a)(b+a)}{2(b-a)}$$
  $$=\frac{a+b}{2}$$

$$\mathrm{Var}(\xi)=\mathbb{E}[\xi^2]-\mathbb{E}[\xi]^2=\dots=\frac{(b-a)^2}{12}$$





Ex Ett godtyckligt tal $a+\xi$ avrundas till a om $\xi \in [-0.5, 0.5]$, där a är ett heltal.
t.ex. $2.8 = 2+0.8 = a+\xi$ och $1.2 = 1+0.2 = a+\xi$.
Avrundningsfelet, $|\xi|$, antas uppfylla att $\xi \sim R[-0.5, 0.5]$. Vad är sannolikheten att hälften av 10 tal har ett avrundningsfel som är större än 0.3 (antas oberoende)?

Lösn

$Y =$ (antal avrundningsfel > 0.3) $\sim Bin(10, p)$

där

$p = P$(ett tal har avrundningsfel > 0.3) $=$

$$= P(|\xi| > 0.3) = \int_{0.3}^{0.5} \frac{1}{0.5-(-0.5)}\,dx + \int_{-0.5}^{-0.3} \frac{1}{0.5-(-0.5)}\,dx = 2\int_{0.3}^{0.5} dx = 2 \cdot 0.2 = 0.4$$

Vi söker

$$P(Y=5) = \binom{10}{5} 0.4^5 \cdot 0.6^5 \approx 0.201$$

Vi har vidare att $E[\xi] = \frac{-0.5+0.5}{2} = 0$

och $\sigma = S(\xi) = \sqrt{Var(\xi)} = \sqrt{\frac{(0.5-(-0.5))^2}{12}} = \sqrt{\frac{1}{12}} \approx 2.89$