Sannolikhet och statistik V

Johan Jonasson

Mars 2019

Fördelning för fkn av sv

Låt X vara en sv och $g : \mathbb{R} \to \mathbb{R}$ en fkn och låt $Y = g(X)$. Vad är fördelningen för Y?

Fallet då X är diskret: $V_X = \{x_1, x_2, \dots\}$.

$$V_Y = \{g(x_j) : x_j \in V_X\}.$$

För varje $y_k \in V_Y$, låt $B_k = \{x_j \in V_X : g(x_j) = y_k\}$. Då gäller

$$p_Y(y_k) = \mathbb{P}(Y = y_k) = \mathbb{P}(X \in B_k) = \sum_{x_j \in B_k} p_x(x_j).$$

Fallet då X är kont: Y kan bli diskret, kontinuerlig eller något annat beroende på hur g ser ut.

Exempel: Låt $X \sim likf[0,1]$ och

$$g(x) = \begin{cases} 0, & x < 2/3 \\ 1, & x \geq 2/3 \end{cases}$$

Med $Y = g(X)$ blir $V_Y = \{0,1\}$ och $p_Y(1) = 1/3$.

Man jobbar oftast från fall till fall.

Exempel. Låt $X \sim likf[0,2]$, dvs $f(x) = 1/2$ och $F(x) = x/2$. Låt

$$Y = X(2-X).$$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X(2-X) \leq y).$$

$$X(2-X) = y \Leftrightarrow X^2 - 2X + y = 0 \Leftrightarrow X = 1 \pm \sqrt{1-y}.$$

Alltså

$$
\begin{aligned}
F_Y(x) = \mathbb{P}(Y \leq y) &= \mathbb{P}(X \leq 1 - \sqrt{1-y}) + \mathbb{P}(X > 1 + \sqrt{1-y}) \\
&= \frac{1-\sqrt{1-y}}{2} + 1 - \frac{1+\sqrt{1-y}}{2} \\
&= 1 - \sqrt{1-y}.
\end{aligned}
$$

Derivering ger också

$$
f_Y(y) = \frac{d}{dy}(1 - \sqrt{1-y}) = \frac{1}{2\sqrt{1-y}}.
$$

Ett generellt resultat: Antag att g är strängt växande. Då finns g^{-1} och är växande. Alltså

$$F_Y(y) = \mathbb{P}(g(X) \leq y) = \mathbb{P}(X \leq g^{-1}(y)) = F(g^{-1}(y)).$$

Alltså

$$f_Y(y) = f_X(g^{-1}(y))(g^{-1})'(y).$$

Exempel: Låt $X \sim \mathit{likf}[0,1]$ och $Y = X^3$. Då är $g(x) = x^3$, så $g^{-1}(y) = y^{1/3}$ och $(g^{-1})'(y) = (1/3)y^{-2/3}$. Eftersom $f(x) = 1$ får vi

$$f_Y(y) = \frac{1}{3}y^{-2/3},\ y \in [0,1].$$

Väntevärde

Låt $X_1, X_2, X_3, \ldots$ vara sv def på samma utfallsrum.

Definition: Man säger att $X_1, X_2, \ldots$ är oberoende om det för alla $B_1, B_2, \ldots \subseteq \mathbb{R}$ gäller att händelserna $\{X_1 \in B_1\}, \{X_2 \in B_2\}, \ldots$ är oberoende.

Låt N vara stort och $X_1, X_2, \ldots, X_N$ vara oberoende och alla ha samma förd som X. Antag att $V_X = \{x_1, x_2, \ldots\}$. Låt $N_k = |\{j : X_j = x_k\}|$. Vi förväntar oss att ha

$$N_k \approx Np(x_k),$$

vilket medför

$$\text{medelvärdet av } X_j\text{:na} = \sum_{k=1}^{\infty} x_k \frac{N_k}{N} \approx \sum_{k=1}^{\infty} x_k p(x_k).$$

(Detta är tyngdpunkten av tallinjen om massorna $p(x_k)$ läggs ut på pos x_k.)

Definition: Väntevärdet av en diskret sv X ges av

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k p(x_k).$$

För en kont sv X gäller att

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

(För def krävs att $\sum_{k=1}^{\infty} |x_k| p(x_k) < \infty$ resp $\int_{-\infty}^{\infty} |x| f(x)dx < \infty$.)

Exempel: Satsa en marker på rött på roulette. Låt X vara vinsten. Då gäller $\mathbb{P}(X = 1) = 18/37$ och $\mathbb{P}(X = -1) = 19/37$. Alltså

$$\mathbb{E}[X] = (-1) \cdot \frac{19}{37} + 1 \cdot \frac{18}{37} = -\frac{1}{37}.$$

Exempel: Vv av diskret likformig: $V_X = \{1, 2, \ldots, n\}$ och $\mathbb{P}(X = k) = 1/n$ för $k = 1, \ldots, n$.

$$\mathbb{E}[X] = \sum_{k=1}^{n} k\frac{1}{n} = \frac{n+1}{2}.$$

Exvis är vv av ett tärningsslag $7/2$.

Exempel: $X \sim likf(a, b)$.

$$\mathbb{E}[X] = \int_a^b x\frac{1}{b-a}dx = \frac{b+a}{2}.$$

Exempel: $X \sim likf(0, 1)$. Låt $Y = X^2$.

$$F_Y(y) = \mathbb{P}(X \leq \sqrt{y}) = \sqrt{y}$$

för $y \in (0, 1)$. Alltså $f_Y(y) = 1/(2\sqrt{y})$.

Alltså

$$\mathbb{E}[Y] = \int_0^1 y \frac{1}{2\sqrt{y}} dy = \frac{1}{3}[y^{3/2}]_0^1 = \frac{1}{3}.$$

Notera att vi alltså har $\mathbb{E}[X^2] = 1/3$ medan $\mathbb{E}[X] = 1/2$ så $\mathbb{E}[X]^2 = 1/4$.

Proposition 2.9: Antag att X är kont och $V_X = [0, \infty)$. Då gäller

$$\mathbb{E}[X] = \int_0^\infty (1 - F(x))dx = \int_0^\infty \mathbb{P}(X > x)dx.$$

Bevis:

$$\begin{aligned}\int_0^\infty \mathbb{P}(X > x)dx &= \int_0^\infty \int_x^\infty f(t)dt\,dx \\ &= \int_0^\infty \left(\int_0^t dx\right) f(t)dt \\ &= \int_0^\infty tf(t)dt = \mathbb{E}[X].\end{aligned}$$

Diskret variant: Antag att $V_X = \{0, 1, 2, \ldots\}$. Då är

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) = \sum_{n=0}^{\infty} \mathbb{P}(X > n).$$

Bevis:

$$\sum_{n=1}^{\infty} \mathbb{P}(X \geq n) = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} p(k) = \sum_{k=1}^{\infty} \sum_{n=1}^{k} p(k) = \sum_{k=1}^{\infty} kp(k).$$

Exempel: I ex. ovan hade vi $F_Y(y) = \sqrt{y}$, $y \in (0, 1)$, så

$$\mathbb{E}[Y] = \int_0^1 (1 - \sqrt{y}) dy = \frac{1}{3}.$$

Exempel: Låt X vara antal tärningskast till första sexan. Vi har

$$\mathbb{P}(X > n) = \left(\frac{5}{6}\right)^n.$$

Alltså

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} \left(\frac{5}{6}\right)^n = \frac{1}{1-\frac{5}{6}} = 6.$$