

Ett sätt att beskriva sannolikhetsfördelningar som fungerar både för diskreta och för kontinuerliga s.v. (och för blandningar av dessa båda sorter av s.v.) är med hjälp av fördelningsfunktionen.

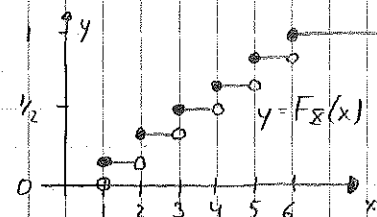
Def: Funktionen $F_X(x) = P(X \leq x)$ kallas fördelningsfunktionen för den s.v. X .

För diskreta s.v. gäller följande samband mellan fördelningsfunktionen och sannolikhetsfunktionen

$$F_X(x) = \sum_{k \leq x} p_X(k) \quad p_X(k) = F_X(k) - F_X(k-1)$$

För kontinuerliga s.v. gäller följande samband mellan fördelningsfunktionen och täthetsfunktionen

$$F_X(x) = \int_{-\infty}^x f_X(t) dt \quad f_X(x) = F'_X(x)$$



Figuren visar fördelningsfunktionen för kast med sexsidig tärning.

Generellt gäller att fördelningsfunktionen har följande egenskaper:

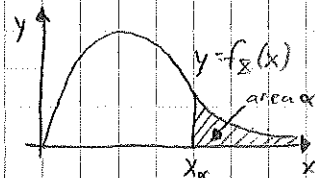
1. $\lim_{x \rightarrow -\infty} F_X(x) = 0$
2. $\lim_{x \rightarrow \infty} F_X(x) = 1$
3. F_X är en icke-avtagande funktion av x .
4. F_X är högerkontinuerlig för alla x .

Sannolikheten att $a < X \leq b$ kan enkelt uttryckas med fördelningsfunktionen. Om $X \leq b$, så sker en av de disjunkta händelserna $X \leq a$ och $a < X \leq b$. Därmed gäller

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

Ofta är det intressant att bestämma sannolikheten för extrema händelser (jfr lab. 0). Det värde på x , för vilket sannolikheten är α att X överstiger x kallas α -kvantilen av X och betecknas x_α . Vi har att

$$P(X > x_\alpha) = \alpha \Rightarrow F_X(x_\alpha) = P(X \leq x_\alpha) = 1 - \alpha.$$



Vi ska nu undersöka några kontinuerliga sannolikhetsfördelningar

Def: En s.v. X sägs vara liktförmigt fördelad på (a, b) om $f_X(x) = \frac{1}{b-a}$ för $x \in (a, b)$ och $f_X(x) = 0$ för $x \notin (a, b)$. Detta betecknas $X \in U(a, b)$. (Om en hållplats passeras av en buss exakt var tredje minut, så är väntetiden $U(0, 10)$ -fördelad.)

Def: En s.v. X sägs vara exponentialfördelad med parameter $\lambda > 0$ om $f_X(x) = \lambda e^{-\lambda x}$ för $x \geq 0$ och $f_X(x) = 0$ annars. Detta betecknas $X \in \text{Exp}(\lambda)$.

En exponentialfördelad s.v. X har fördelningsfunktionen

$$F_X(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^x = -e^{-\lambda x} + 1$$

Exponentialfunktionen används ofta för att modellera väntetider. Antag att en affär passeras av n personer per tidsenhet och att de förbi-passerade går in i affären med sannolikhet p . Låt X vara tiden till första kunden, så $X > x$ motsvarar att ingen tid kommit vid tiden x . Låt $\mu = np$, där n är stort, p litet och μ har storleksordning ett.

$$P(X > x) = (1-p)^{nx} = \left(1 - \frac{\mu}{n}\right)^n \approx (e^{-\mu/n})^n = e^{-\mu x} \Rightarrow F_X(x) = 1 - e^{-\mu x} \Rightarrow X \in \text{Exp}(\mu)$$

personer som passerat vid tiden x

Exponentialfördelningen har en intressant minnesförlustegenskap. Låt X beteckna livslängden hos en komponent (= "vänteliden" tills den går sönder). Sannolikheten att en komponent som överlevt t tidsenheter överlever ytterligare x tidsenheter är

$$\begin{aligned} P(X > t+x | X > t) &= \frac{P(X > t+x, X > t)}{P(X > t)} = \frac{P(X > t+x)}{P(X > t)} = \frac{1 - P(X \leq t+x)}{1 - P(X \leq t)} \\ &= \frac{1 - F_X(t+x)}{1 - F_X(t)} = \frac{1 - (1 - e^{-\lambda(t+x)})}{1 - (1 - e^{-\lambda t})} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = 1 - F_X(x) = P(X > x), \end{aligned}$$

så den kvarvarande livslängden är oberoende av komponentens ålder.

Def: En s.v. X sägs vara normalfördelad (eller Gaussisk) om $f_X(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}$, där $\mu \in \mathbb{R}$, $\sigma > 0$. Detta betecknas $X \in N(\mu, \sigma)$.

Normalfördelningen är den viktigaste sannolikhetsfördelningen, främst eftersom summor av s.v. ofta är approximativt normalfördelade oavsett fördelningen hos de s.v. Detta kallas centrala gränsvärdesatsen (se FL 7).

Om X är en s.v. (diskret eller kontinuerlig) och g är en reellvärd funktion, så är $Y = g(X)$ också en s.v. Exempel på detta är $Y = X^2$, $Y = \sqrt{X}$ eller $Y = aX + b$. Om g är strikt monoton, så kan fördelningen för $Y = g(X)$ bestämmas med följande metod

Ex: Låt $X \in U(0,1)$ och $g(x) = -\frac{1}{\lambda} \ln x$ för något $\lambda > 0$. Det gäller att:

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) = P\left(-\frac{1}{\lambda} \ln X \leq y\right) = P(\ln X \geq -\lambda y) \\ &= P(X \geq e^{-\lambda y}) = 1 - P(X < e^{-\lambda y}) = 1 - F_X(e^{-\lambda y}) \end{aligned}$$

$P(X = e^{-\lambda y}) = 0$, ty X är en kont. s.v.

I allmänhet gäller $F_Y(y) = 1 - F_X(g^{-1}(y))$ om g är en strikt avtagande funktion och $F_Y(y) = F_X(g^{-1}(y))$ om g är en strikt växande funktion. $F_X(x)$ är känd (se höger), så vi får

$$F_Y(y) = 1 - F_X(e^{-\lambda y}) = \begin{cases} 1 - e^{-\lambda y}, & y \geq 0 \\ 0, & y < 0 \end{cases} \Rightarrow Y \in \text{Exp}(\lambda) \quad \left| \quad F_X(x) = \begin{cases} 0, & x \leq 0 \\ x, & x \in (0,1) \\ 1, & x \geq 1 \end{cases} \right.$$

Transformationen g ger en metod för att generera exponentialfördelade slumpstal från likförmigt fördelade slumpstal.

Ofte är det vid ett slumpförsök av intresse att registrera flera olika storheter och dessa kan beskrivas med en n -dimensionell s.v. $(X_1, X_2, \dots, X_n)$. Vi kommer här att nöja oss med det tvådimensionella fallet (och skiljer även här mellan diskreta och kontinuerliga s.v.).

Def: En tvådimensionell s.v. (X, Y) är en funktion från Ω till $\mathbb{R}^2$.

Def: En tvådimensionell s.v. (X, Y) sägs vara diskret om såväl X som Y antar ett ändligt eller uppräkneligt oändligt antal värden. Den simultana sannolikhetsfunktionen för en diskret tvådimensionell s.v. (X, Y) ges av $p_{X,Y}(j,k) = P(X=j, Y=k)$ för tillåtna värden på j och k . Notera att $0 \leq p_{X,Y}(j,k) \leq 1$, $\sum_j \sum_k p_{X,Y}(j,k) = 1$ och $P((X,Y) \in A) = \sum_{(j,k) \in A} p_{X,Y}(j,k)$.

Def: En tvådimensionell s.v. (X, Y) sägs vara kontinuerlig om det existerar en icke-negativ funktion $f_{X,Y}$, kallad den simultana täthetsfunktionen för (X, Y) , sådan att

$$P((X, Y) \in A) = \iint_A f_{X,Y}(x,y) dx dy \quad \text{för alla } A \in \mathbb{R}^2 \quad \left(\iint_{\mathbb{R}^2} f_{X,Y}(x,y) dx dy = 1 \right)$$