Föreläsning 5

Särskilda kontinuerliga fördelningar

Definition (a) En *stokastisk variabel* X är en avbildning (funktion)

$$X : S \longrightarrow \mathbb{R} .$$

(b) En stokastisk variabel X sägs vara *kontinuerlig* om den kan anta alla värden av ett intervall eller union av intervall.

(c) För en kontinuerlig stokastisk variabel definieras *täthetsfunktionen* $f_X : \mathbb{R} \longrightarrow [0, \infty)$ genom

$$P(a < X \leq b) = \int_a^b f_X(x)\, dx\,, \quad -\infty < a \leq b < \infty.$$

(d)

$$F_X(x) = P(X \leq x)\,, \quad x \in \mathbb{R},$$

kallas *fördelningsfunktion*.

Exempel på kontinuerliga stokastisk variabler

- X: livslängd av en glödlampa
- Y: vikt av en regndroppe

Egenskaper

(1) $f_X \geq 0$ (täthetsfunktion är icke-negativ)

(2) Hela sannolikhetsmassan är 1

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

(3)

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b)$$

$$P(X = a) = 0 \quad \text{om} \quad X \text{ är kontinuerlig}$$

Obs. Detta stämmer inte om X är diskret.

Exempel Antag att X är en stokastisk variabel med täthetsfunktion

$$f_X(x) = \begin{cases} c \cdot (4x - 2x^2) & \text{om} \quad 0 < x < 2 \\ 0 & \text{annars} \end{cases} .$$

(a) Hitta c.
(b) Bestäm $P(X > 1)$.

Svar: (a)

$$1 = \int_{-\infty}^{\infty} f_X(x)\,dx = c \cdot \int_{-\infty}^{\infty} (4x - 2x^2)\,dx$$
$$= c \cdot \left(2x^2 - \frac{2}{3}x^3\right)\Big|_0^2 = c \cdot \frac{8}{3}\,.$$

Detta medför $c = \dfrac{3}{8}$.

(b)

$$P(X > 1) = \int_1^2 f_X(x)\,dx = c \cdot \int_1^2 (4x - 2x^2)\,dx$$
$$= c \cdot \left(2x^2 - \frac{2}{3}x^3\right)\Big|_1^2 = \cdots = \frac{1}{2}\,.$$

Väntevärde och varians

Definition Låt X vara en kontinuerlig stokastisk variabel. (a)

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx$$

kallas *väntevärde* av X.

(b)

$$Var(X) \equiv V(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)\,dx$$

kallas *varians* av X.

(c)

$$D(X) = \sqrt{Var(X)}$$

kallas *standardavvikelse* av X.

Anmärkningar (1) Vi säger att väntevärdet *existerar* om högra sidan är absolut integrerbar, dvs.

$$\int_{-\infty}^{\infty} |x| \cdot f_X(x)\,dx < \infty.$$

(2) Ofta fördelaktigt att använda den så kallade *förkortningsformeln*,

$$Var(X) = E[X^2] - (E[X])^2$$

$$= \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\,dx - (E[X])^2\,.$$

Sats (Räkneregler) (a) $E[a] = a$, $Var(a) = 0$ för varje icke-slumpmässigt $a \in \mathbb{R}$.

(b)

$$E[aX + b] = aE[X] + b\,, \quad a, b \in \mathbb{R},$$

$$Var(aX+b) = a^2 Var(X)\,, \quad a, b \in \mathbb{R}, \quad \text{(inget } b \text{ å högra sidan).}$$

(c) Låt g vara en reell funktion. Det gäller att

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx$$

om högra sidan är absolut integrerbar, dvs. $\int_{-\infty}^{\infty} |g(x)| \cdot f_X(x)\, dx < \infty$.

Anmärkningar (1) Med (c) bevisas andra delen av förkortningsformeln. Här tas $g(x) = x^2$, $x \in \mathbb{R}$.
(2) Vi säger att variansen och standardavvikelsen *existerar* om

$$\int_{-\infty}^{\infty} x^2 \cdot f_X(x)\, dx < \infty.$$

Detta medför att väntevärdet existerar.

Exempel Låt X vara en kontinuerlig stokastisk variabel med täthetsfunktion

$$f_X(x) = \begin{cases} 2x & \text{om} \quad 0 \le x \le 1 \\ 0 & \text{annars} \end{cases} .$$

Beräkna $E[X]$ och $Var(X)$.

Svar:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx = \int_0^1 x \cdot 2x\, dx = \cdots = \frac{2}{3}.$$

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \\ &= \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\, dx - (E[X])^2 \\ &= \int_0^1 x^2 \cdot 2x\, dx - \left(\frac{2}{3}\right)^2 = \cdots = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}\,. \end{aligned}$$

Exempel Låt X vara en kontinuerlig stokastisk variabel med täthetsfunktion

$$f_X(x) = \begin{cases} 1 & \text{om} \quad 0 \le x \le 1 \\ 0 & \text{annars} \end{cases} .$$

Beräkna $E[e^X]$ och $Var(e^X)$.

Svar:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx \quad \text{med} \quad g(x) = e^x\,.$$

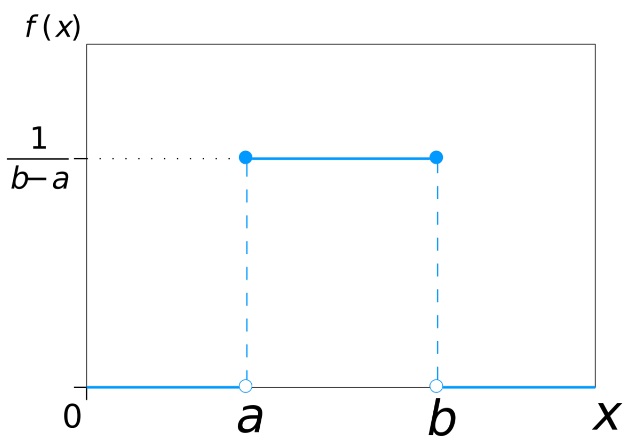
$$E\left[e^X\right] = \int_0^1 e^x \cdot 1\, dx = e^x|_0^1 = e - 1.$$

$$
\begin{aligned}
Var\left(e^X\right) &= E\left[\left(e^X\right)^2\right] - \left(E\left[e^X\right]\right)^2 \\
&= E\left[e^{2X}\right] - (e-1)^2 = \int_0^1 e^{2x}\cdot 1\,dx - (e-1)^2 \\
&= \frac{1}{2}\,e^{2x}\Big|_0^1 - (e-1)^2 = \cdots = -\frac{1}{2}e^2 + 2e - \frac{3}{2}\,.
\end{aligned}
$$

Särskilda kontinuerliga fördelningar

Definition En kontinuerlig stokastisk variabel sägs vara *likformigt fördelad* på intervallet (a, b) eller $[a, b]$ ($U(a, b)$-fördelad, $Re(a, b)$-fördelad) om

$$
f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{om} \quad a < x < b \\ 0 & \text{annars} \end{cases}.
$$



Google Bild

Exempel Vid en busshållplats kommer det bussar var 12:e minut. En person kommer till hållplatsen vid en slumpmässig tidpunkt. Låt X vara hennes / hans väntetid tills en buss kommer. Bestäm sannolikheten att personen väntar längre än 9 minuter.

Svar: $X \sim U(0, 12)$. Sökt är $P(X > 9)$.

$$
P(X > 9) = \int_9^{12} f_X(x)\,dx = \int_9^{12} \frac{1}{12}\,dx = \frac{1}{4}\,.
$$

Sats Låt $X \sim U(a, b)$. Det gäller att

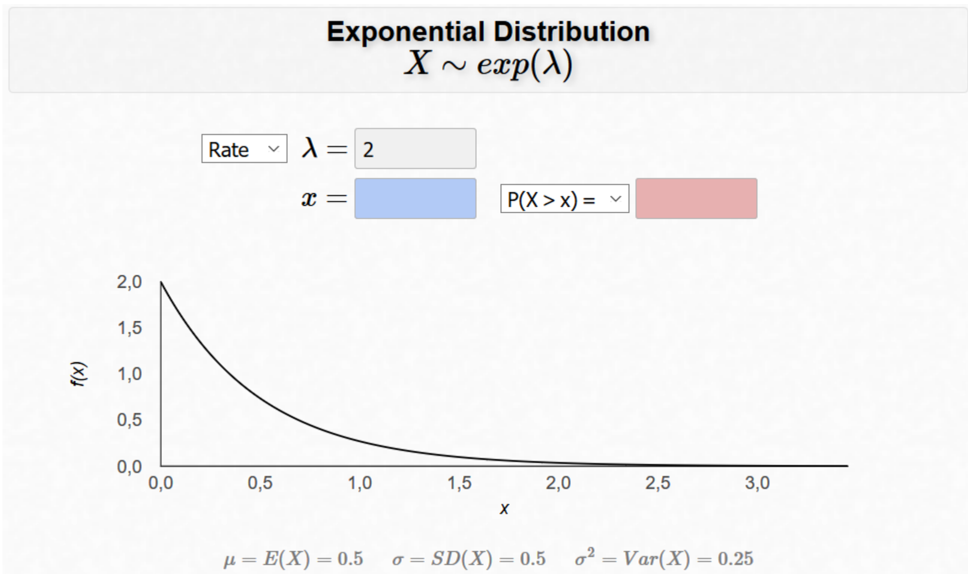
$$
F_X(x) = \begin{cases} 0 & \text{om} \ -\infty < x \le a \\ \dfrac{x-a}{b-a} & \text{om} \quad a < x < b \\ 1 & \text{om} \quad b \le x < \infty \end{cases}
$$

och

$$
E[X] = \frac{a+b}{2} \quad \text{och} \quad Var(X) = \frac{(b-a)^2}{12}.
$$

Definition En kontinuerlig stokastisk variabel sägs vara *exponentialfördelad* med parametern $\lambda > 0$ ($Exp(\lambda)$-fördelad) om

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{om} \quad x \geq 0 \\ 0 & \text{för övrigt} \end{cases} .$$



Matt Bognar U Iowa

Sats Låt $X \sim Exp(\lambda)$. Det gäller att

$$F_X(x) = \begin{cases} 0 & \text{om} \quad -\infty < x < 0 \\ 1 - e^{-\lambda x} & \text{om} \quad 0 \leq x < \infty \end{cases} ,$$

$$E[X] = \frac{1}{\lambda} \quad \text{och} \quad Var(X) = \frac{1}{\lambda^2}.$$

Exempel Man har observerat tiderna mellan successiva utbrott av vulkanen *Mauna Loa* på Havaii och kommit fram att dessa tider är exponentialfördelade med det skattade väntevärdet 36.722 månader. Man observerar vulkanen ytterligare 80 månader. Vad är sannolikheten att ett utbrott upplevas?

Svar: Låt X vara tid mellan två successiva utbrott. Det gäller att

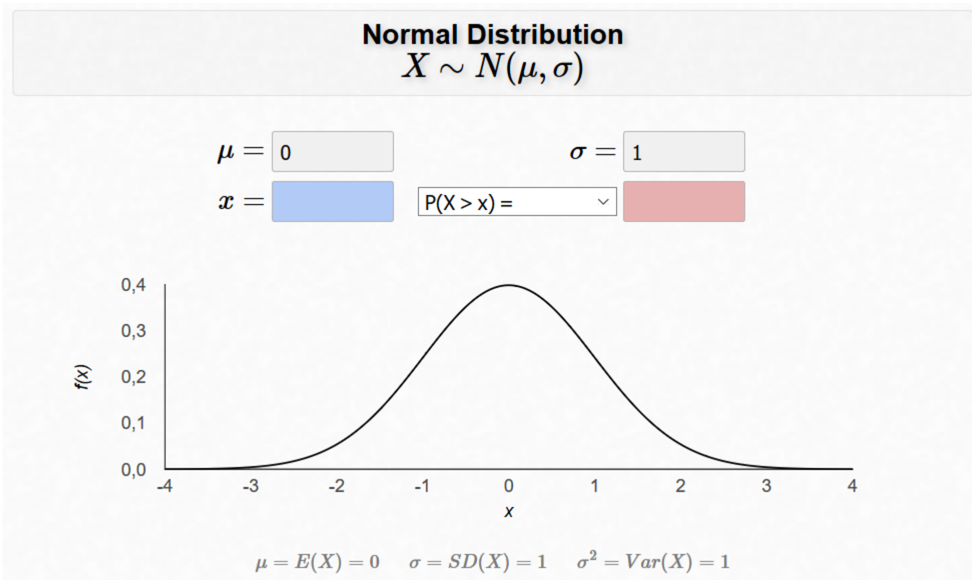
$$E[X] = \frac{1}{\lambda} = 36.722 \quad \text{som ger} \quad \lambda = \frac{1}{36.722} .$$

Sökt är

$$P(X \leq 80) = F_X(80) = 1 - e^{-(1/36.722)\cdot 80} = 0.887 .$$

Definition En kontinuerlig stokastisk variabel Z sägs vara *standard normalfördelad* ($N(0,1)$-fördelad) om

$$\varphi_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} , \quad x \in \mathbb{R}.$$



Matt Bognar U Iowa

Anmärkningar (1) En $N(0,1)$-fördelad stokastisk variabel betecknas ofta med Z. Täthetsfunktionen för en $N(0,1)$-fördelad stokastisk variabel betecknas ofta med $\varphi \equiv \varphi_Z$. Fördelningsfunktionen betecknas ofta med Φ, dvs.

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^{x} \varphi_Z(y)\, dy\,, \quad x \in \mathbb{R}.$$

(2) Integralen

$$\int_a^b \varphi_Z(x)\, dx$$

kan inte algebraiskt beräknas. För att bestämma $P(a < Z \leq b)$ använder man $N(0,1)$-tabellen. Hur man använder $N(0,1)$-tabellen förklaras under denna länk.

Sats Om $Z \sim N(0,1)$ så är

$$E[Z] = 0 \quad \text{och} \quad Var(Z) = 1\,.$$

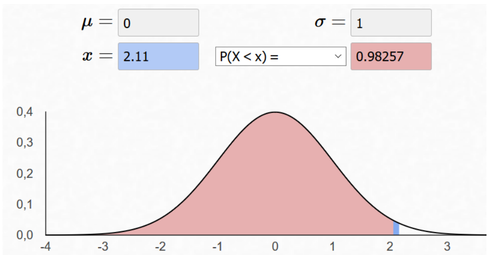
Exempel Låt $Z \sim N(0,1)$. Bestäm (a) $P(Z < 2.11)$.
(b) $P(Z \leq -2.21)$.
(c) $P(-2.21 < Z < 2.11)$.

Ledning: (1) Rita upp en figur. (2) Använd komplement och symmetri.

Svar:

(a) Direkt ur tabellen fås

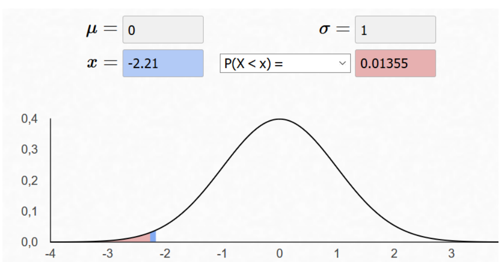
$$P(Z < 2.11) = \Phi(2.11) = 0.9826\,.$$



Matt Bognar U Iowa

(b) Använd symmetri och komplement för att få

$$P(Z \le -2.21) = \Phi(-2.21) = 1 - \Phi(2.21) = 1 - 0.9864 = 0.0136\,.$$



Matt Bognar U Iowa

(c) Använd resultaten från (a) och (b) för att få

$$P(-2.21 < Z < 2.11) = P(Z < 2.11) - P(Z \le -2.21) = 0.9826 - 0.0136 = 0.969\,.$$

Definition En kontinuerlig s.v. sägs vara *normalfördelad* med parametrarna μ och σ ($N(\mu, \sigma)$-fördelad) om

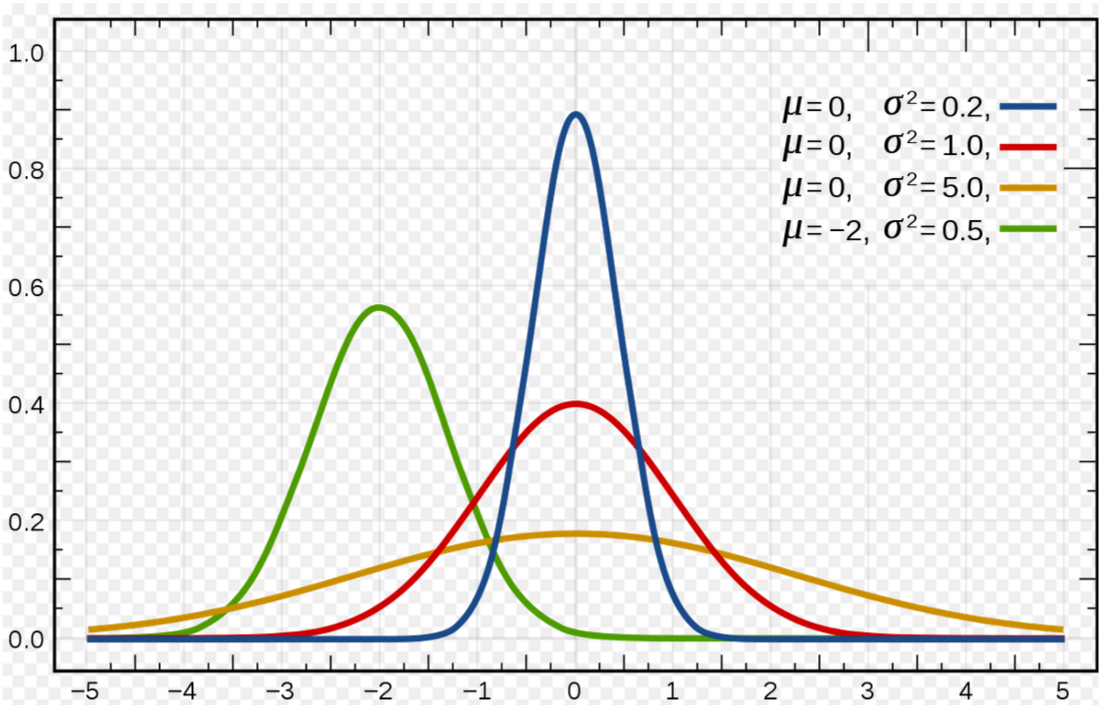
$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,, \quad x \in \mathbb{R}.$$

Anmärkning Integralen

$$\int_a^b f_X(x)\,dx$$

kan inte algebraiskt beräknas. För att bestämma $P(a < Z \le b)$ använder man standard-iseringen och $N(0,1)$-tabellen.

Täthetsfunktion $N(\mu, \sigma)$ för olika μ, σ.



Wikipedia

Sats Låt $X \sim N(\mu, \sigma)$. (a) Det gäller att

$$E[X] = \mu \quad \text{och} \quad Var(X) = \sigma^2.$$

(b) (Standardisering)

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

och

$$P\left(a < X \leq b\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Exempel Låt $X \sim N(3, 3)$. Bestäm $P(2 < X < 5)$.

Svar:

$$\begin{aligned} P(2 < x < 5) &= \Phi\left(\frac{5-3}{3}\right) - \Phi\left(\frac{2-3}{3}\right) \\ &= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) = \Phi\left(\frac{2}{3}\right) - \left(1 - \Phi\left(\frac{1}{3}\right)\right) \\ &= 0.7486 - (1 - 0.6293) = 0.3779\,. \end{aligned}$$

Jörg-Uwe Löbus, jorg-uwe.lobus@liu.se