Sannolikhetslära och Statistik
Föreläsning 4

Mattias Villani

Avdelningen för Statistik och Maskininlärning
Institutionen för datavetenskap
Linköpings universitet

Översikt

- Täthetsfunktion
- Likformig fördelning
- Exponentialfördelningen
- Gammafördelningen
- Normalfördelningen

Kontinuerliga slumpvariabler

- Kontinuerliga slumpvariabler kan anta alla reela värden på ett inteval (a, b), speciellt $(-\infty, \infty)$.
- X kontinuerlig $\Rightarrow P(x) = 0$ för alla x. Pmf inte användbar.
- Fördelningsfunktionen funkar dock: $F(x) = \mathbf{P}\{X \leq x\}$.
- Eftersom $P(x) = 0$ för alla x så gäller $\mathbf{P}\{X \leq x\} = \mathbf{P}\{X < x\}$.
- Om X kontinuerlig slumpvariabel: $F(x)$ **kontinuerlig**. Inga hopp. **Icke-avtagande**.

$$\lim_{x \to \infty} F(x) = 1 \qquad \lim_{x \to -\infty} F(x) = 0.$$

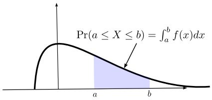
Täthetsfunktion

Definition. **Täthetsfunktionen** $f(x)$ för en kontinuerlig slumpvariabel X är derivatan av CDF:n

$$f(x) = F'(x).$$

- Fördelningen är kontinuerlig om den har en täthetsfunktion.
- Täthetsfunktion heter **probability density function**, **pdf** på engelska.
- cdf:n $F(x)$ är antiderivatan av pdf:n.
- Sannolikheter för intervall ges av ytor under pdf:n

$$\mathbf{P}\{a < X < b\} = \int_a^b f(x)dx$$



Täthetsfunktion

- $f(x) = F'(x)$ så

$$\int_{-\infty}^{b} f(x)dx = F(b) - F(-\infty) = F(b) - 0 = F(b).$$

- Täthetsfunktioner integrerar till ett:

$$\int_{-\infty}^{\infty} f(x)dx = F(\infty) - F(-\infty) = 1 - 0 = 1.$$

- Täthetsfunktionens värden, t ex $f(2)$, är inte en sannolikhet. $f(2) > 1$ helt ok. Men $f(x) \geq 0$ måste gälla.
- För litet ϵ: $\Pr\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) \approx \epsilon \cdot f(a)$.
- `Exempel`: triangelfördelningen över support $[0, a]$. Normaliseringskonstant. Fördelningsfunktion. $P\{X > a/2\}$. Se också Example 4.1 i Baron.
- Se Table 4.1 i Baron för en jämförelse av diskreta och kontinuerliga fördelningar.

Väntevärde och varians

- Repetition: för diskreta slumpvariabler:

$$\mathbb{E}X = \sum_x x \cdot P(x) \quad Var(X) = \mathbb{E}\left(X-\mu\right)^2 = \sum_x \left(x-\mu\right)^2 P(x)$$

- För kontinuerliga slumpvariabler:

$$\mathbb{E}X = \int x \cdot f(x)dx \qquad Var(X) = \mathbb{E}\left(X-\mu\right)^2 = \int (x-\mu)^2 f(x)dx$$

- Exempel: triangelfördelning.

Simultanfördelning för kontinuerliga variabler

- **Simultan fördelningsfunktion**

$$F_{(X,Y)}(x,y) = \mathrm{P}\left\{X \leq x \cap Y \leq y\right\}$$

- **Simultan täthetsfunktion**

$$f_{(X,Y)}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x,y)$$

- Ofta skriver vi bara $f(x,y)$ istället för $f_{(X,Y)}(x,y)$.
- **Kovarians**

$$Cov(X,Y) = \mathbb{E}\left(X-\mu_X\right)\left(Y-\mu_Y\right) = \int\int \left(X-\mu_X\right)\left(Y-\mu_Y\right) f(x,y$$

Likformig fördelning

- **Täthetsfunktion** för likformig fördelad slumpvariabel över support $[a, b]$

$$f(x) = \frac{1}{b-a} \quad \text{för } a \leq x \leq b,$$

och $f(x) = 0$ annars.

- Väntevärde

$$\mathbb{E}X = \int x \cdot f(x)dx = \frac{1}{b-a}\int xdx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b = \frac{1}{2(b-a)}\left(b^2 -\right.$$

- Varians

$$Var(X) = \mathbb{E}X^2 - \mu^2$$

$$\mathbb{E}X^2 =$$