NMAI059 Probability and statistics 1
Class 9

Robert Šámal

Overview

Continuous random vectors

Covariance and correlation

Inequalities

Limit theorems – approximation

What we know

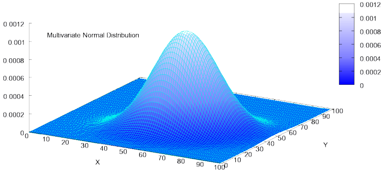
- joint cdf

$$F_{X,Y}(x, y) = P(X \leq x \,\&\, Y \leq y).$$

- joint pdf: $f_{X,Y} \geq 0$ such that

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t) dt ds.$$

- important example: multivariate normal distribution



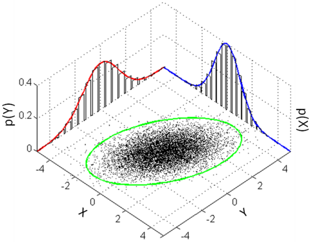


Image by Wikipedia editors Piotrg and Bscan.

Conditioning

Definition (restricting a r.v. to a subset)

X is a r.v. on $(\Omega, \mathcal{F}, P)$, $B \in \mathcal{F}$, s.t. $P(B) > 0$.

$$F_{X|B}(x) := P(X \leq x \mid B)$$

$f_{X|B}$ is the corresponding pdf.

- if $B = \{X \in S\}$, then

$$f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P(X \in S)} & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

Total cdf & pdf

Theorem (total cdf, total pdf)

Let X be a continuous r.v., let B_1, B_2, $\ldots$ be a partition of Ω. Then

$$F_X(x) = \sum_i P(B_i) F_{X|B_i}(x) \quad \textit{and}$$

$$f_X(x) = \sum_i P(B_i) f_{X|B_i}(x).$$

Proof: law of total probability.

Marginal pdf

Theorem

$$f_X(x) = \int_{y \in \mathbb{R}} f_{X,Y}(x,y)dy$$

$$f_Y(y) = \int_{x \in \mathbb{R}} f_{X,Y}(x,y)dx$$

Conditional pdf

Definition

For continuous r.v. X, Y we define their conditional pdf by

$$f_{X|Y}(x|y) := \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

when $f_Y(y) > 0$, otherwise we do not define it.

- recall that $f_Y(y) = \int_{x \in \mathbb{R}} f_{X,Y}(x,y)dx$
- for a fixed y the function $x \mapsto f_{X|Y}(x|y)$ is a pdf

Conditional, joint and marginal pdf

Theorem

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y)$$
$$f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y) dy$$

Sum of continuous r.v.

Theorem

Let X, Y be independent random variables. Then $Z = X + Y$ is also a continuous r.v. and its pdf is a convolution of f_X, f_Y. That is,

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx.$$

Example of a convolution

Conditional density and expectation

- $\mathbb{E}\big(X \mid B\big) := \int_{-\infty}^{\infty} x \cdot f_{X|B}(x)dx$
- $\mathbb{E}\big(g(X)\big|B\big) = \int_{-\infty}^{\infty} g(x)f_{X|B}(x)dx$

Theorem (total expectation)

Let X be a continuous r.v. If B_1, B_2, $\ldots$ is a partition of Ω, then

$$\mathbb{E}\big(X\big) = \sum_i P(B_i)\mathbb{E}\big(X \mid B_i\big).$$

Proof: by total pdf.

Conditional pdf and expectation

- $f_{X|Y}(x|y) := \frac{f_{X,Y}(x,y)}{f_Y(y)}$ is a pdf of X, given $Y = y$
- $\mathbb{E}(X \mid Y = y) := \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x, y)dx$ is the expectation of this r.v.
- $\mathbb{E}(g(X)|Y = y) = \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y}(x, y)dx$
- An analogy of the law of total expectation:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} \mathbb{E}(X \mid Y = y) f_Y(y)dy$$

- $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid Y))$

Overview

Covariance

Definition

For r.v.'s X, Y we define their covariance by formula

$$cov(X,Y) = \mathbb{E}\big((X - \mathbb{E}X)(Y - \mathbb{E}Y)\big).$$

Theorem

$$cov(X,Y) = \mathbb{E}\big(XY\big) - \mathbb{E}\big(X\big)\mathbb{E}\big(Y\big)$$

- $var(X) = cov(X,X)$
- $cov(X, aY + bZ + c) = a\,cov(X,Y) + b\,cov(X,Z)$
- $cov(X,Y) = 0$ if X,Y are independent
- but not only then

Correlation

Definition

Correlation of random variables X, Y is defined by

$$\varrho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)\, var(Y)}}.$$

- "scaled covariance"
- $-1 \leq \varrho(X, Y) \leq 1$ (exercise)
- Correlation does not imply causation! (In particular, correlation is symmetric.)
- OTOH, uncorrelation does not imply independence. (Extreme case: X any r.v., $Y = +X$ or $Y = -X$, both with the same probability.)

Variance of a sum

Theorem

Let $X = \sum_{i=1}^{n} X_i$*. Then*

$$var(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} cov(X_i, X_j) = \sum_{i=1}^{n} var(X_i) + \sum_{i \neq j} cov(X_i, X_j).$$

In particular, if $X_1, \ldots, X_n$ *are independent, then*

$$var(X) = \sum_{i=1}^{n} var(X_i).$$

Overview

Continuous random vectors

Covariance and correlation

Inequalities

Limit theorems – approximation

Cauchy inequality

Theorem

Let X, Y have finite expectation and variance. Then

$$\mathbb{E}(XY) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$$

- Corollary for correlation: $-1 \leq \varrho(X,Y) \leq 1$

Jensen inequality

Theorem

Let X have finite expectation and let g be a convex real functin. Then

$$\mathbb{E}\big(g(X)\big) \geq g(\mathbb{E}\big(X\big)).$$

(For concave function we have the opposite inequality.)

Markov inequality

Theorem

Suppose $X \geq 0$ *and* $a > 0$*. Then*

$$P(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Chebyshev inequality

Theorem

Let X have finite expectation μ and variance σ^2, let $a > 0$. Then

$$P(|X - \mu| \geq a \cdot \sigma) \leq \frac{1}{a^2}.$$

Chernoff inequality

Theorem

Let $X = \sum_{i=1}^{n} X_i$, where X_i are i.i.d. attaining ± 1 with probability $1/2$. Then for $t > 0$ we have

$$P(X \leq -t) = P(X \geq t) \leq e^{-t^2/2\sigma^2},$$

where $\sigma = \sigma_X = \sqrt{n}$.

Without proof.

Overview

Strong law of large numbers

Theorem

Let $X_1, \ldots, X_n$ be i.i.d. with expectation μ and variance σ^2. Let $S_n = (X_1 + \cdots + X_n)/n$ be the sample mean. Then we have

$$\lim_{n\to\infty} S_n = \mu \quad \textit{almost surely (i.e. with probability 1).}$$

We say that sequence S_n converges to μ almost surely.

Monte Carlo integration

How to compute $\int_{x \in A} g(x) dx$?
In particular

$$g(x) = \begin{cases} 1 & \text{for } x \in S \\ 0 & \text{otherwise} \end{cases}$$

... area of a circle

Weak law of large numbers

Theorem

Let $X_1, \ldots, X_n$ be i.i.d. with expectation μ and variance σ^2. Let $S_n = (X_1 + \cdots + X_n)/n$ be the sample mean. Then for every $\varepsilon > 0$ we have

$$\lim_{n \to \infty} P(|S_n - \mu| > \varepsilon) = 0.$$

We say that sequence S_n converges to μ in probability.

Central Limit Theorem

Central Limit Theorem

Theorem

Let $X_1, \ldots, X_n$ *be i.i.d. with expectation* μ *and variance* σ^2*. Put* $Y_n := ((X_1 + \cdots + X_n) - n\mu)/(\sqrt{n} \cdot \sigma)$*.*
Then $Y_n \xrightarrow{d} N(0,1)$*. This means, that if* F_n *is the cdf of* Y_n*, then*

$$\lim_{n\to\infty} F_n(x) = \Phi(x) \quad \textit{for every } x \in \mathbb{R}.$$

We say that the sequence Y_n *converges to* $N(0,1)$ *in distribution.*

Moment generating function

Definition

For a random variable X we let

$$M_X(t) = \mathbb{E}\left(e^{tX}\right).$$

Function $M_X(t)$ is called the moment generating function.

- $M_{Bern(p)}(t) = p \cdot e^t + (1-p)$.
- $M_X(t) = \sum_{n=0}^{\infty} \mathbb{E}\left(X^n\right) \frac{t^n}{n!}$.
- $M_{X+Y}(t) = M_X(t) M_Y(t)$, jsou-li X, Y n.n.v.
- $M_{Bin(n,p)} = (pe^t + 1 - p)^n$
- $M_{N(0,1)} = e^{t^2/2}$
- $M_{Exp(\lambda)} = \frac{1}{1-t/\lambda}$
- If $M_X(t) = M_Y(t)$ on $(-a, a)$ for some $a > 0$, then $X = Y$ a.s.