Review - Random vectors II: correlation, convolution

Covariance and correlation:
The covariance $\operatorname{cov}(X, Y)$ of random variables X, Y is defined as

$$\operatorname{cov}(X, Y) = \mathrm{E}(X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y) = \mathrm{E}(XY) - (\mathrm{E}\,X)(\mathrm{E}\,Y),$$

if $\mathrm{E}\,X^2 < \infty$ and $\mathrm{E}\,Y^2 < \infty$.
We compute

$$\mathrm{E}\,XY = \sum_{ij} x_i y_j \,\mathrm{P}(X = x_i \cap Y = y_j) \qquad \text{for discrete } X, Y,$$

and

$$\mathrm{E}\,XY = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\,\mathrm{d}x\,\mathrm{d}y \qquad \text{for continuous } X, Y.$$

The correlation coefficient ρ_{XY} is defined as

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var} X}\sqrt{\operatorname{var} Y}}$$

if $\operatorname{var} X, \operatorname{var} Y > 0$. Necessarily $-1 \le \rho_{XY} \le 1$. The correlation coefficient represents a measure of linear dependence between X and Y.

Note: If X and Y are independent, then $\operatorname{cov}(X, Y) = 0$. The reverse does not always hold.

Further properties: If $X_1, \ldots, X_n$ are random variables, then (if exists):

- $\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{E}\,X_i,$
- $\operatorname{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{var} X_i + \sum\sum_{i \ne j} \operatorname{cov}(X_i, X_j).$

Sums of two random variables (convolution):

a) When X, Y are independent with a discrete distribution, then the random variable $Z = X + Y$ has a discrete distribution with probablities

$$\mathrm{P}(Z = n) = \sum_{k=-\infty}^{\infty} \mathrm{P}(X = k)\,\mathrm{P}(Y = n - k).$$

b) When X, Y are independent with a continuous distribution with densities f_X and f_Y, then the random variable $Z = X + Y$ has a continuous distribution with density

$$g(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,\mathrm{d}x.$$

Minimum and maximum:

a) If X and Y are independent, then

$$\mathrm{P}(\min(X, Y) > z) = \mathrm{P}(X > z \cap Y > z) = \mathrm{P}(X > z)\,\mathrm{P}(Y > z),$$
$$\mathrm{P}(\max(X, Y) \le z) = \mathrm{P}(X \le z \cap Y \le z) = \mathrm{P}(X \le z)\,\mathrm{P}(Y \le z).$$

b) For any X and Y it holds that:

$$\min(X, Y) + \max(X, Y) = X + Y.$$

Exercises 7 - Random vectors II: correlation, convolution

1. Variance of a sum

a) Show that for two non-correlated random variables X and Y it holds that

$$\operatorname{var}(X+Y) = \operatorname{var} X + \operatorname{var} Y.$$

b) Try to find a formula for the variance if X and Y are correlated.

c) Which of the previous formulas holds if X and Y are independent?

2. Let X_1 and X_2 be independent random variables with common expectation μ and variance σ^2.

a) Find the expected values and the variances of the random variables

$$Y_1 = X_1 + X_2, \qquad Y_2 = 2X_1, \qquad Y_3 = X_1 - X_2.$$

b) Compute the covariance $\operatorname{cov}(Y_i, Y_j)$ for $i, j = 1, 2, 3$, with $i \neq j$.

c) What do we know about the independence or the pairwise independence of the random variables Y_1, Y_2, Y_3?

3. Let X and Y be two independent random variables with geometric distribution with parameter p.

a) Find the distribution of $\min\{X, Y\}$.
Hint: *Compute* $\mathrm{P}(\min\{X, Y\} > k)$.

b) Find the expected value of $\max\{X, Y\}$.
Hint: $X + Y = \min\{X, Y\} + \max\{X, Y\}$.

c) Find the distribution of $\max\{X, Y\}$.
Hint: *Compute* $\mathrm{P}(\max\{X, Y\} \leq k)$.

d) Compute $\mathrm{P}(X < Y)$.
Hint: *Sum up the probabilities of favorable cases over all possible values of* X.

4. Let $X_1, X_2, \ldots, X_n$ be independent random variables. Denote

$$Y = \min\{X_1, \ldots, X_n\}.$$

a) Suppose that all X_i have a discrete uniform distribution on $\{1, 2, \ldots, k\}$. Find the distribution of the random variable Y.

b) Suppose that all X_i have a geometric distribution with parameter p. Find the distribution of the random variable Y.

5. Let $X_1, X_2, \ldots, X_n$ be independent random variables. Denote

$$Y = \min\{X_1, \ldots, X_n\}.$$

Suppose that each X_i is exponentially distributed with parameter λ_i.

a) Find the distribution of the random variable Y. (Begin with $n = 2$)
Hint: *Compute* $\mathrm{P}(Y > k)$.

b) Find the expected value of $\max\{X_1, X_2\}$.
Hint: $X_1 + X_2 = \min\{X_1, X_2\} + \max\{X_1, X_2\}$.

6. Suppose that for two random variables X and Y it holds that:

$$\mathrm{E}\, X = 1, \qquad \mathrm{E}\, XY = 2.$$

Furthermore, suppose that the random variable Y is a linear function of X (i.e., $Y = aX + b$) and the variables X and Y are non-correlated. Find the variance of Y.

7. Consider two random variables X and Y and constants $a > 0$, b, $c > 0$, d. Prove that for the correlation coefficient it holds that:

$$\rho(aX + b, cY + d) = \rho(X, Y).$$

Additional exercises - Random vectors II: correlation, convolution

Random vectors continued

8. Suppose that X and Y are independent random variables with Poisson distribution with parameters λ_1 and λ_2 respectively. Find the distribution of $Z = X + Y$.

9. Suppose that X and Y are two independent random variables with exponential distribution with parameters λ_1 and λ_2 respectively. Find the distribution of $Z = \min(X, Y)$
Hint: *use the function* $S_Z(z) = \mathrm{P}(\min(X, Y) > z)$.

10. Suppose that $X_1, \ldots, X_n$ are independent and identically distributed (iid) continuous random variables with the same distribution function F and density f. Find the distribution function and density of $Z = \max(X_1, \ldots, X_n)$.

11. Suppose that $X_1, \ldots, X_n$ are independent and identically distributed (iid) continuous random variables with $\mathrm{E}\, X_1 = \mu$ and $\operatorname{var} X_1 = \sigma^2$. Denote $S_n = \sum_{i=1}^{n} X_i$. What is the expectation and variance of S_n?

12. Let X have an uniform distribution on the interval $[-1, 1]$. Denote $Y = X^2$. What are the covariance and the coefficient of correlation between X and Y? Are they independent?

13. * Suppose there are n gentlemen at a theatre, each putting their hat in the cloak room. On their way home, each of them gets one hat assigned at random. Consider random variables $X_1, \ldots, X_n$ with $X_i = 1$ if the i^{th} gentleman has his own hat and $X_i = 0$ if he doesn't.

a) Find the expectation and variance of X_i for a given i.

b) Are X_i and X_j independent for $i \neq j$? What is the covariance of X_i and X_j?

c) What is the mean and variance of the total number of correctly assigned hats $X = \sum_{i=1}^{n} X_i$?

14. Saint Martin is approaching, and in your restaurant you're preparing for a busy day. Let W the number of bottles of young wine drunk and G be the random variable describing the number of geese eaten, in one minute. Suppose that W and G are continuous with the joint density

$$f_{W,G}(w, g) = \begin{cases} \frac{1}{2} e^{-w-g/2} & \text{for } g > 0,\ w > 0, \\ 0 & \text{otherwise.} \end{cases}$$

a) What are the marginal distributions of W and G?

b) Are W and G independent?

c) After how much time are you expecting the first wine bottle to be emptied? And after how much time the first goose to be finished?

d) What is the covariance of W and G?

e) Wht is the total expected time $E(W + G)$?

