5. Vector random variables, functions of random variables

Vector random variables

Joint probability: joint probability distribution of $\boldsymbol{Z} = (X, Y)$:

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = F_{XY}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x, y)\, dx\, dy,$$

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}.$$

Marginal probability distribution of X:

$$F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy, \qquad f_X(x) = \frac{\partial F_X(x)}{\partial x} = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy.$$

Conditional probability density of X given $Y = y$:

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

Independent random variables X and Y:

$$f_{X|Y}(x|y) = f_X(x) \iff f_{XY}(x, y) = f_X(x) \cdot f_Y(y).$$

Functions of random variables

Function of a scalar variable: In case $f_X(x)$ and the function $Y = g(X)$ are known, $f_Y(y)$ can be determined using the inverse function h, defined by $X = h(Y) = g^{-1}(Y)$:

$$f_Y(y) = f_X\big(h(y)\big) \left|\frac{dh(y)}{dy}\right|.$$

The above equation is valid for *monotonous* $g(X)$. If $g(X)$ is not monotonous, its domain is divided into k *piecewise monotonous* parts $g_i(X)$ with the corresponding inverse functions $h_i(Y)$:

$$f_Y(y) = \sum_{i=1}^{k} f_X\big(h_i(y)\big) \left|\frac{dh_i(y)}{dy}\right|.$$

Scalar function of a vector variable: In case $f_{XY}(x, y)$ and the function $Z = g(X, Y)$ are known, $f_Z(z)$ can be determined. The formulation of $f_Z(z)$ in general depends on the function $g(X, Y)$. In the simplest case, where $Z = g(X, Y) = X + Y$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(x, z - x)\, dx \xRightarrow{X,Y \text{ independent}} \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\, dx.$$

If X and Y are *randomly independent*, the result is the *convolution integral* on the right-hand side.

Sum (difference) of two independent normal random variables X and Y: probability distribution of $Z = X \pm Y$ is determined by the convolution integral:

$$\mathcal{N}(X; m_X, \sigma_X), \mathcal{N}(Y; m_Y, \sigma_Y) \xRightarrow{Z=X\pm Y} \mathcal{N}\left(Z; m_Z = m_X \pm m_Y, \sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2}\right).$$

The mean m_Z is a sum (difference) of means, while the variance σ_Z^2 is always a sum of variances.

5. Vector random variables, functions of random variables - problems

1. (a) Determine the value of c such that the function $f_{XY}(x, y) = cxy$ for $0 < x < 3$ and $0 < y < 3$ satisfies the properties of a joint probability density function. Determine the following: (b) $P(X < 2, Y < 3)$, (c) $P(1 < Y < 2.5)$, (d) $f_Y(y)$, (e) $f_{Y|X}(y|1.5)$. R: (a) 4/81, (b) 4/9, (c) 0.583, (d) $2y/9$, (e) $2y/9$

2. Random variable X has a probability distribution $f_X(x) = 2x/9$ for $0 \leq x < 3$. Find the probability distribution of the random variable $Y = 2X + 5$. R: $f_Y(y) = (y - 5)/18$

3. Lifetime of two light bulbs is exponentially distributed with averages of $1/a$ and $1/b$, respectively, where $a \neq b$ and $a, b \neq 0$. Light bulbs are connected in such a way that only one is in use. When the first one burns out the second lights up. Determine the probability density function of the total time of lighting. R: $f_Z(z) = ab(a - b)^{-1}(e^{-bz} - e^{-az})$

4. Shaft diameter at the point of compression joint is normally distributed with a mean of 100 mm and standard deviation 0.2 mm. The diameter of a hole in a gear is normally distributed with a mean of 99 mm and the same standard deviation as that of the shaft diameter. The compression joint is of a good quality if the shaft diameter is 0.5 to 2.5 mm larger than the hole diameter. What is the probability that a randomly selected gear does not make a good quality joint with a randomly selected shaft? R: P = 0.038

5. Pieces with volume of 0.6 m^3 are painted by dipping in a container with volume of 1 m^3. For each piece, the paint is poured in the container by two machines. The amount of the poured paint is normally distributed for both of them. The mean paint volume poured by the first machine is 0.25 m^3, the standard deviation is 0.03 m^3, while the other machine pours paint with mean of 0.1 m^3 and the standard deviation 0.01 m^3. After painting each piece, the remaining paint is removed from the container. What is the probability that during the painting of a randomly selected piece the paint does not flow over the container edge? R: P = 0.943

6. It is known that in drilling of mild steel the total length of the holes that can be drilled by one drill bit in its lifetime is normally distributed with a mean of 1 m and a standard deviation of 0.2 m. At least how many drill bits are needed to drill a 6 m long hole with a probability of 95 %? R: n = 7

7. We have 10000 m^3 of gas in stock. It is assumed that the daily gas consumption is normally distributed with an average of 50 m^3 and a standard deviation of 5 m^3. It is also assumed that the consumption of a randomly selected day is independent of the consumption in other days. For how many days will our gas stock last, if there has to be enough gas with a probability of at least 0.95? R: n = 197

5. Vector random variables, functions of random variables – additional problems

1. (a) Determine the value of c that makes the function $f_{XY}(x,y) = c(x+y)$ a joint probability density function over the range $0 < x < 3$ and $x < y < x+2$. Determine the following: (b) $P(X < 1, Y < 2)$, (c) $P(Y > 1)$, (d) $f_X(x)$, (e) $f_{X|Y}(x|2)$. R: (a) 1/24, (b) 5/48, (c) 143/144, (d) $(2x+1)/12$, (e) $(x+2)/6$ for $x < 2$

2. A random variable X has the probability density function $f_X(x) = x/8$ for $0 \leq x < 4$. Determine the probability density function of a variable $Y = 2X + 4$. R: $f_Y(y) = (y-4)/32$

3. A random variable X has the probability density function:

$$f_X = \begin{cases} \dfrac{C}{1+(x-1)^2}, & \text{for } 0 \leq x < 2, \\ 0, & \text{for } x \text{ elsewhere.} \end{cases}$$

   Determine the probability density function of variable $Y = (X-1)^2$. R: $f_Y(y) = 2/[\pi\sqrt{y}(1+y)]$

4. Lifetime of a machine component is exponentially distributed with an average of $\frac{1}{\lambda}$. When a component fails, it is replaced with a spare, which has the same characteristics as the original component and both are independent of each other. The machine works as long as one of the two components work. Determine the probability density function of the machine lifetime. R: $f_Z(z) = \lambda^2 z\, e^{-\lambda z}$

5. In assembly of the product a piece having a hole has to be joined with a piece having a plug. Depth of the hole and height of the plug are normally distributed. Average depth of the hole is 10.50 mm with a standard deviation of 0.20 mm, while the plug has an average height of 10.38 mm with a standard deviation of 0.18 mm.

   a. What is the probability that a randomly selected plug is too high for a randomly selected hole? R: P = 0.328

   b. What should the average depth of hole at a given standard deviation and at given plug distribution be to have the probability of the first problem case less than 0.05? R: m_l > 10.82 mm

6. Daily consumption of water in the village is limited to 1000 m^3. Daily water consumption of households is normally distributed with an average of 600 m^3 and standard deviation of 70 m^3. Daily water consumption of industry is also normally distributed with an average of 200 m^3 and a standard deviation of 40 m^3. In addition to the household and industry consumption, water is also used for other purposes. What is the maximal amount of water that can be daily used for other purposes so that the village limit consumption is not exceeded with a probability of at least 0.95? R: V < 67.4 m^3