

Methods: Statistics (4,120)

# 6. Multi-Dimensional Probability Distributions

Spring 2022

**Prof. Dr. Roland Füss**

Swiss Institute of Banking and Finance (s/bf)

# Contents

- I. Joint Probability Function**
- II. Joint Cumulative Distribution Function**
- III. Marginal Distributions**
- IV. Stochastic Independence**
- V. Conditional Distributions**
- VI. Expected Value and Variance**
- VII. Covariance under Stochastic Independence**

## Learning Objectives

After this lecture, you know how:

- the probability concepts for **two-dimensional random variables** are applied in theory and in practice.
- the concepts of covariance and correlation are applied to random variables and how they are used to confirm **stochastic independence**.

## Literature

Shira, Joseph (2012). *Statistische Methoden der VWL und BWL*, 4th ed., Munich et al.: Pearson Studium, **Chapter 10**.

# I. Joint Probability Function

The **joint probability function**  $f(x_i, y_k)$  states the probability that  $X$  takes the value  $x_i$  and at the same time  $Y$  takes the value  $y_k$ .  $X, Y$  are **discrete** random variables with realizations  $x_1, x_2, \dots$  and  $y_1, y_2, \dots$

$$f(x_i, y_k) = P(X = x_i, Y = y_k)$$

$$f(x_i, y_k) \geq 0$$

$$\sum_i \sum_k f(x_i, y_k) = 1$$

# I. Joint Probability Function

**Machine breakdowns:** A machine has sensitive components B1 and B2.  $X$  is the number of breakdowns per day from B1,  $x_1 = 0, x_2 = 1, x_3 = 2$ .  $Y$  is the number of breakdowns per day of B2,  $y_1 = 0, y_2 = 1, y_3 = 2$ . Thus, up to 2 breakdowns per day are modeled. The joint probability function  $f(x_i, y_k)$  is given by:

$$f(x_i, y_k) = P(X = x_i, Y = y_k)$$

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.30	0.14	0.02
$X = 1$	0.18	0.10	0.02
$X = 2$	0.12	0.06	0.06

## II. Joint Cumulative Distribution Function

The **joint cumulative distribution function**  $F(x, y)$  states the probability that  $X$  takes at most the value  $x$  and at the same time  $Y$  takes at most the value  $y$ .  $X, Y$  are again discrete random variables.

$$F(x, y) = \text{Prob}(X \leq x, Y \leq y)$$

Relationship between distribution function and probability function:

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_k \leq y} f(x_i, y_k)$$

## II. Joint Cumulative Distribution Function

**Machine breakdowns:**

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.30	0.44	0.46
$X = 1$	0.48	0.72	0.76
$X = 2$	0.60	0.90	1

The joint cumulative distribution function follows from the relationship between distribution and probability function:

$$F(x, y) = \sum_{x_i \leq x} \sum_{y_k \leq y} f(x_i, y_k)$$

### III. Marginal Distributions (Example)

Machine breakdowns:

	$Y = 0$	$Y = 1$	$Y = 2$	$f_x(x)$
$X = 0$	0.30	0.14	0.02	0.46
$X = 1$	0.18	0.10	0.02	0.30
$X = 2$	0.12	0.06	0.06	0.24
$f_y(y)$	0.60	0.30	0.10	1

$$f_X(x_i) = P(X = x_i) = \sum_k f(x_i, y_k)$$

$$f_Y(y_k) = P(Y = y_k) = \sum_i f(x_i, y_k)$$

## IV. Stochastic Independence

The **independence** of two discrete random variables  $X$ ,  $Y$  is given if the following condition is fulfilled:

$$f(x_i, y_k) = f_x(x_i) \cdot f_y(y_k)$$

**Analogous:** Multiplication rule for independent events:

$$P(A \cap B) = P(A)P(B)$$

The **marginal distribution**  $f_x(x_i)$  for  $X$  states the probability for  $X = x_i$  no matter what value  $Y$  accepts (in the case of stochastic independence):

$$f_x(x_i) = P(X = x_i) = \sum_k f(x_i, y_k)$$

## V. Conditional Distributions

### Machine breakdowns:

The conditional distribution of  $X$  given  $Y$  is defined by :

$$f(x_i|y_k) = \frac{f(x_i, y_k)}{f_y(y_k)}$$

Assume  $Y = 0$  occurred:

$$f(x_i | Y = 0) = \frac{f(x_i, 0)}{f_Y(0)}$$

$X$	$f(x_i, 0)$	$f(x_i Y = 0)$
$i = 0$	0.30	0.50
$i = 1$	0.18	0.30
$i = 2$	0.12	0.20
$f_y(0)$	0.6	

## V. Conditional Distributions

### Machine breakdowns:

The conditional distribution of  $Y$  given  $X$  is defined by :

$$f(y_k|x_i) = \frac{f(x_i, y_k)}{f_x(x_i)}$$

Assume  $X = 2$  occurred:

$$f(y_k | X=2) = \frac{f(2, y_k)}{f_x(2)}$$

$X$	$f(2, y_k)$	$f(y_k X = 2)$
$k = 0$	0.12	0.50
$k = 1$	0.06	0.25
$k = 2$	0.06	0.25
$f_x(2)$	0.24	

## VI. Expected Value of Marginal Distributions

### Machine breakdowns:

The expected values  $E(X)$  and  $E(Y)$  can be calculated via marginal distributions:

	$Y = 0$	$Y = 1$	$Y = 2$	$f_x(x)$
$X = 0$	0.30	0.14	0.02	0.46
$X = 1$	0.18	0.10	0.02	0.30
$X = 2$	0.12	0.06	0.06	0.24
$f_y(y)$	0.60	0.30	0.10	1

$$E(X) = \sum_{i=1}^3 x_i f_X(x_i) = 0 \cdot 0.46 + 1 \cdot 0.30 + 2 \cdot 0.24 = 0.78$$

$$E(Y) = \sum_{k=1}^3 y_k f_Y(y_k) = 0 \cdot 0.60 + 1 \cdot 0.30 + 2 \cdot 0.10 = 0.50$$

## VI. Variance

### Machine breakdowns:

The variances  $Var(X)$  and  $Var(Y)$  can be determined via the marginal distributions and the expected values  $E(X)$  and  $E(Y)$ :

$$\begin{aligned} Var(X) &= \sum_{i=1}^3 x_i^2 f_X(x_i) - [E(X)]^2 \\ &= 0 \cdot 0.46 + 1 \cdot 0.30 + 4 \cdot 0.24 - 0.78^2 \\ &= 0.6516 \end{aligned}$$

$$\begin{aligned} Var(Y) &= \sum_{i=1}^3 y_k^2 f_Y(y_k) - [E(Y)]^2 \\ &= 0 \cdot 0.60 + 1 \cdot 0.30 + 4 \cdot 0.10 - 0.5^2 \\ &= 0.45 \end{aligned}$$

## VI. Conditional Expected Value and Variance

### Machine breakdowns:

The conditional expected values as well as the conditional probabilities can be determined (assume  $Y = 0$  occurred):

$X$	$f(x_i, 0)$	$f(x_i   Y = 0)$
$i = 0$	0.30	0.50
$i = 1$	0.18	0.30
$i = 2$	0.12	0.20
$f_y(0)$	0.60	

$$E(X | Y = 0) = \sum_{i=1}^3 x_i f(x_i | Y = 0) = 0 \cdot 0.50 + 1 \cdot 0.30 + 2 \cdot 0.20 = 0.70$$

$$\begin{aligned} \text{Var}(X | Y = 0) &= \sum_{i=1}^3 x_i^2 f(x_i | Y = 0) - [E(X | Y = 0)]^2 \\ &= 0 \cdot 0.50 + 1 \cdot 0.30 + 4 \cdot 0.20 - 0.70^2 = 0.61 \end{aligned}$$

## VI. Conditional Expected Value and Variance

### Machine breakdowns:

The conditional expected values as well as the conditional probabilities can be determined (assume  $X = 2$  occurred):

$Y$	$f(2, y_k)$	$f(y_k   X = 2)$
$k = 0$	0.12	0.50
$k = 1$	0.06	0.25
$k = 2$	0.06	0.25
$f_x(2)$	0.24	

$$E(Y | X = 2) = \sum_{k=1}^3 y_k f(y_k | X = 2) = 0 \cdot 0.50 + 1 \cdot 0.25 + 2 \cdot 0.25 = 0.75$$

$$\begin{aligned} \text{Var}(Y | X = 2) &= \sum_{k=1}^3 y_k^2 f(y_k | X = 2) - [E(Y | X = 2)]^2 \\ &= 0 \cdot 0.50 + 1 \cdot 0.25 + 4 \cdot 0.25 - 0.75^2 = 0.6875 \end{aligned}$$

## VI. Expected Value and Variance (R-Example 1)

Open the file "L6-Example\_1.R" in R-Studio and reproduce the R-Code.

```
# there are two discrete random variables X and Y with the following possible
# realizations  $S_X = 0, 5, 10$  and  $S_Y = 0, 5, 10, 15$ . the common
# probability function of X and Y is given by the matrix p.
p <- matrix(c(.02,.04,.01,.06,.15,.15,.02,.20,.14,.10,.10,.01),ncol=4)
sum(p) # all probabilities should add up to one

# what is the probability P that X=5 and Y=10?
p[2,3]

# marginal distributions
#-----
# calculate the marginal distributions for x.
px <- apply(p,1,sum)

# calculate the marginal distributions for y.
py <- apply(p,2,sum)

# conditional distributions
#-----
# calculation of the conditional probability  $P(X=5|Y=5)$ .
p_x5_y5 <- p[2,2]/py[2]

# calculate the conditional probability  $P(X=0|Y=5)$  and  $P(X=10|Y=5)$ .
p_x0_y5 <- p[1,2]/py[2]
p_x10_y5 <- p[3,2]/py[2]

# calculation of the conditional probability  $P(X|Y=5)$ .
p_x_y5 <- c(p_x0_y5,p_x5_y5,p_x10_y5)

# expected value and variance
#-----
# realizations of the two random variables
x<- c(0,5,10)
y<- c(0,5,10,15)

# berechnung des Erwartungswertes  $E(X)$  und  $E(X^2)$ .
EX <- sum(px*x)
EX2 <- sum(px*x^2)
```

## VII. Covariance and Correlation Coefficient

**Covariance** and **correlation coefficient** are measures for the strength of a statistical relationship between two random variables. If  $Y$  tends to large values whenever  $X$  takes large values, then  $X$  and  $Y$  are **(positively) correlated**.

$$\begin{aligned}\text{Cov}(X, Y) &= \sum_i \sum_k [x_i - E(X)] \cdot [y_k - E(Y)] \cdot f(x_i, y_k) \\ &= E(XY) - E(X) \cdot E(Y)\end{aligned}$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \cdot \sqrt{\text{Var}(Y)}} \in [-1, +1]$$

## VII. Covariance and Correlation Coefficient

$$\begin{aligned}\text{Cov}(X, Y) &= \sum_i \sum_k [x_i - E(X)] \cdot [y_k - E(Y)] \cdot f(x_i, y_k) \\ &= \sum_i \sum_k x_i y_k f(x_i, y_k) - E(X) \cdot E(Y) \\ &= E(XY) - E(X) \cdot E(Y)\end{aligned}$$

$$\begin{aligned}E(XY) &= \sum_i \sum_k x_i y_k f(x_i, y_k) \\ &= 0 \cdot 0 \cdot 0.30 + 0 \cdot 1 \cdot 0.14 + 0 \cdot 2 \cdot 0.02 \\ &\quad + 1 \cdot 0 \cdot 0.18 + 1 \cdot 1 \cdot 0.10 + 1 \cdot 2 \cdot 0.02 \\ &\quad + 2 \cdot 0 \cdot 0.12 + 2 \cdot 1 \cdot 0.06 + 2 \cdot 2 \cdot 0.06 \\ &= 0.50\end{aligned}$$

$$\begin{aligned}\text{Cov}(X, Y) &= E(XY) - E(X) \cdot E(Y) \\ &= 0.50 - 0.78 \cdot 0.50 = 0.11\end{aligned}$$

$$\rho = \frac{0.11}{\sqrt{0.6516} \cdot \sqrt{0.45}} = 0.20$$

## VII. Covariance under Stochastic Independence

Stochastic Independence:  $f(x_i, y_k) = f_x(x_i) \cdot f_y(y_k)$

$$\begin{aligned} E(XY) &= \sum_i \sum_k x_i y_k f(x_i, y_k) \\ &= \sum_i \sum_k x_i y_k f_x(x_i) f_y(y_k) \\ &= \sum_i x_i f_x(x_i) \sum_k y_k f_y(y_k) \\ &= E(X)E(Y) \end{aligned}$$

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY) - E(X) \cdot E(Y) \\ &= E(X) \cdot E(Y) - E(X) \cdot E(Y) = 0 \end{aligned}$$

## VII. Covariance under Stochastic Independence

### Summary:

$$\begin{aligned}\text{independence} &\Rightarrow \text{Cov}, \rho = 0 \\ \text{Cov}, \rho = 0 &\not\Rightarrow \text{independence}\end{aligned}$$

or (in general):

$$\begin{aligned}\text{Cov}, \rho \neq 0 &\Rightarrow \text{dependence} \\ \text{Cov}, \rho = 0 &\not\Rightarrow \text{independence}\end{aligned}$$

## Appendix: Covariance and Correlation Coefficient

The **coefficient of correlation**  $\rho$  always takes values between -1 and +1:

$$\rho(X, Y) \in [-1, +1]$$

Proof:

$$\text{Define the RV : } E[((X - E[X]) - \lambda(Y - E[Y]))^2] \geq 0 \quad \forall \lambda$$

$$E[((X - E[X]) - \lambda(Y - E[Y]))^2] \geq 0 \quad \forall \lambda$$

$$E[(X - E[X])^2 - 2\lambda(X - E[X])(Y - E[Y]) + \lambda^2(Y - E[Y])^2]$$

$$= \text{Var}(X) - 2\lambda \text{Cov}(X, Y) + \lambda^2 \text{Var}(Y) \geq 0 \quad \forall \lambda$$

$$\lambda := \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}$$

$$\text{Var}(X) - \frac{\text{Cov}^2(X, Y)}{\text{Var}(Y)} \geq 0$$

$$1 \geq \frac{\text{Cov}^2(X, Y)}{\text{Var}(Y) \cdot \text{Var}(X)} = \rho(X, Y)^2$$