# Advanced Mathematical Methods

## WS 2023/24

## 4 Mathematical Statistics

Dr. Julie Schnaitmann

*Department of Statistics, Econometrics and Empirical Economics*

Eberhard Karls Universität Tübingen

Wirtschafts- und Sozialwissenschaftliche Fakultät

# Outline: Mathematical Statistics

# Readings

- A. Papoulis and A. U. Pillai. *Probability, Random Variables and Stochastic Processes*.
  Mc Graw Hill, fourth edition, 2002, Chapters 1-4

# Online References

MIT Course on Probabilistic Systems Analysis and Applied Probability (by John Tsitsiklis)

- Discrete RVs I: Concept of random variables, probability mass function, expected value, variance
  https://www.youtube.com/watch?v=3MOahpLxj6A
- Continuous RVs: probability density function, cumulative distribution function, expected value, variance
  https://www.youtube.com/watch?v=mHfn_7ym6to
- Discrete RVs II: Functions of RV, conditional probabilities, specific distribution, total expectation theorem, joint probabilities
  https://www.youtube.com/watch?v=-qCEoqpwjf4

# 4.1 Random Variables

A random variable $X$ takes on real numbers according to some distribution.

There are two types of random variables:

1. discrete random variables
   - e.g. coin toss, number of baskets scored out of $n$ trials
2. continuous random variables
   - e.g. financial returns

# 4.1 Random Variables

A random variable $X$ takes on real numbers according to some distribution.

There are two types of random variables:

1. discrete random variables
    - e.g. coin toss, number of baskets scored out of $n$ trials
2. continuous random variables
    - e.g. financial returns

# 4.1 Random Variables

## Random sample

$\{X_1, X_2, \ldots, X_n\}$ is called a random sample if

1. all draws $X_i$ are **independent**
2. and drawn from the same distribution, i.e. they are **identically distributed**

$\Rightarrow$ the draws are **independently and identically distributed** in short **iid**

# 4.2 Cumulative Distribution Functions

## Probability distribution function: discrete case

$$f_X(x_i) = P(X = x_i)$$

requirements:

- $0 \leq P(X = x_i) \leq 1$
- $\sum_{x_i} f_X(x_i) = 1$

# 4.2 Cumulative Distribution Functions

## (Probability) Density function: continuous case

$f_X(x)$ is not a probability as $P(X = x) = 0$

<u>requirements:</u>

- $P(a \leq X \leq b) = \int_a^b f_X(x) \, \mathrm{d}x \geq 0$
- $\int_{-\infty}^{\infty} f_X(x) \, \mathrm{d}x = 1$

# 4.2 Cumulative Distribution Functions

**Definition: Cumulative distribution function**

The cumulative distribution function (cdf) of a random variable $X$ is defined to be the function $F_X(x) = P(X \leq x)$, for $x \in \mathbb{R}$.

**discrete:**

$$F_X(x_i) = \sum_{X \leq x_i} f_X(x_i) = P(X \leq x_i)$$

**continuous:**

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$$

# 4.2 Cumulative Distribution Functions

## Properties

1) $F_X(+\infty) = 1;\ F_X(-\infty) = 0$

2) $F_X(x)$ is a nondecreasing function of $x$:
if $x_1 < x_2$, $F_X(x_1) \leq F_X(x_2)$
note: the event $\{X \leq x_1\}$ is a subset of $\{X \leq x_2\}$

3) if $F_X(x_0) = 0$, then $F_X(x) = 0 \quad \forall \quad x \leq x_0$

# 4.2 Cumulative Distribution Functions

## Properties

1) $F_X(+\infty) = 1$; $F_X(-\infty) = 0$

2) $F_X(x)$ is a nondecreasing function of $x$:
   if $x_1 < x_2$, $F_X(x_1) \leq F_X(x_2)$
   note: the event $\{X \leq x_1\}$ is a subset of $\{X \leq x_2\}$

3) if $F_X(x_0) = 0$, then $F_X(x) = 0 \quad \forall \quad x \leq x_0$

# 4.2 Cumulative Distribution Functions

## Properties

1) $F_X(+\infty) = 1$; $F_X(-\infty) = 0$

2) $F_X(x)$ is a nondecreasing function of $x$:
   if $x_1 < x_2$, $F_X(x_1) \leq F_X(x_2)$
   note: the event $\{X \leq x_1\}$ is a subset of $\{X \leq x_2\}$

3) if $F_X(x_0) = 0$, then $F_X(x) = 0 \quad \forall \quad x \leq x_0$

# 4.2 Cumulative Distribution Functions

## Properties

4) $P(X > x) = 1 - F_X(x)$
   events $\{X \leq x\}$ and $\{X > x\}$ are mutually exclusive and $\{X \leq x\} \cup \{X > x\} = \Omega$

5) $F_X(x)$ is continuous from the right:
   $\lim_{x \to a^+} F_X(x) = F_X(a)$

6) $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$

# 4.2 Cumulative Distribution Functions

## Properties

4) $P(X > x) = 1 - F_X(x)$
events $\{X \leq x\}$ and $\{X > x\}$ are mutually exclusive and $\{X \leq x\} \cup \{X > x\} = \Omega$

5) $F_X(x)$ is continuous from the right:
$\lim_{x \to a^+} F_X(x) = F_X(a)$

6) $P(x_1 \leq X \leq x_2) = F_X(x_2) - F_X(x_1)$

# 4.2 Cumulative Distribution Functions

## Properties

4) $P(X > x) = 1 - F_X(x)$
events $\{X \le x\}$ and $\{X > x\}$ are mutually exclusive and $\{X \le x\} \cup \{X > x\} = \Omega$

5) $F_X(x)$ is continuous from the right:
$\lim_{x \to a^+} F_X(x) = F_X(a)$

6) $P(x_1 \le X \le x_2) = F_X(x_2) - F_X(x_1)$

# 4.3 Expectation, Variance and Moments

## Expectations of a random variable

$$E[X] = \begin{cases} \sum\limits_{x_i} x_i f_X(x_i) & \text{if } x \text{ is discrete} \\ \int\limits_{-\infty}^{\infty} x f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

If $g(X)$ a measurable function of $x$, then:

$$E[g(X)] = \begin{cases} \sum\limits_{x_i} g(x_i) f_X(x_i) & \text{if } x \text{ is discrete} \\ \int\limits_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

# 4.3 Expectation, Variance and Moments

## Expectations of a random variable

$$E[X] = \begin{cases} \sum_{x_i} x_i f_X(x_i) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} x f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

If $g(X)$ a measurable function of $x$, then:

$$E[g(X)] = \begin{cases} \sum_{x_i} g(x_i) f_X(x_i) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $E[a] = a$
- $E[bX] = b \cdot E[X]$
- linear transformation: $E[a + bX] = a + bE[X]$
- $E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

# 4.3 Expectation, Variance and Moments

## Variance of a random variable

Let $g(X) = (X - E[X])^2$

$$Var[X] = \sigma^2 = E[(X - E[X])^2]$$

$$= \begin{cases} \sum_{x_i} (x_i - E[X])^2 f_X(x_i) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x) \mathrm{d}x & \text{if } x \text{ is continuous} \end{cases}$$

## 4.3 Expectation, Variance and Moments

### Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

**important result:**

$Var[X] = E[X^2] - E[X]^2$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

**important result:**

$Var[X] = E[X^2] - E[X]^2$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

**important result:**

$$Var[X] = E[X^2] - E[X]^2$$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

**important result:**

$Var[X] = E[X^2] - E[X]^2$

# 4.3 Expectation, Variance and Moments

## Calculation rules

- $Var[a] = 0$
- $Var[X + a] = Var[X]$
- $Var[bX] = b^2 Var[X]$
- $Var[a + bX] = b^2 Var[X]$

**important result:**

$Var[X] = E[X^2] - E[X]^2$

# 4.3 Expectation, Variance and Moments

Standardization of a random variable $X$

Let

$$g(X) = \frac{X - \mu}{\sigma} = Z$$

$$Z = \frac{X - \mu}{\sigma} = \frac{-\mu}{\sigma} + \frac{1}{\sigma} X$$

$\Rightarrow \quad E[Z] = 0 \qquad$ and $\qquad Var[Z] = 1$

# 4.3 Expectation, Variance and Moments

Standardization of a random variable $X$

Let

$$g(X) = \frac{X - \mu}{\sigma} = Z$$

$$Z = \frac{X - \mu}{\sigma} = \frac{-\mu}{\sigma} + \frac{1}{\sigma}X$$

$\Rightarrow \quad E[Z] = 0 \qquad$ and $\qquad Var[Z] = 1$

# 4.3 Expectation, Variance and Moments

## Chebychev Inequality

For any random variable $X$ with finite expected value $\mu$ and finite variance $\sigma^2 > 0$ and a positive constant $k$

$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

# 4.3 Expectation, Variance and Moments

## Skewness and Kurtosis

Central moments of a random variable:

$$\mu_r = E[(X - \mu)^r]$$

as $r$ grows, $\mu_r$ tends to explode

Solution: normalization

- skewness coefficient: $\gamma = \dfrac{E[(X-\mu)^3]}{\sigma^3}$
- kurtosis: $\kappa = \dfrac{E[(X-\mu)^4]}{\sigma^4}$

  often reported as excess kurtosis $\kappa - 3$

# 4.3 Expectation, Variance and Moments

## Skewness and Kurtosis

Central moments of a random variable:

$$\mu_r = E[(X-\mu)^r]$$

as $r$ grows, $\mu_r$ tends to explode

Solution: normalization

- skewness coefficient: $\gamma = \dfrac{E[(X-\mu)^3]}{\sigma^3}$
- kurtosis: $\kappa = \dfrac{E[(X-\mu)^4]}{\sigma^4}$
  often reported as excess kurtosis $\kappa - 3$

# 4.4 Quantile

**Quantile**

$q\%$ of the probability mass of a random variable is left of $x(q)$ .

Example: Risk measure Value-at-risk (VaR)

$$q = P(X \leq x(q)) = F(x(q))$$

# 4.5 Specific probability distributions

## The normal distribution

$X$ is a Gaussian or normal random variable with parameters $\mu$ and $\sigma^2$ if its density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

denoted $X \sim N(\mu, \sigma^2)$

**Linear transformation is also normally distributed:**

If $X \sim N(\mu, \sigma^2)$, then $a + bX \sim N(a + b\mu, b^2\sigma^2)$.

# 4.5 Specific probability distributions

## The normal distribution

$X$ is a Gaussian or normal random variable with parameters $\mu$ and $\sigma^2$ if its density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

denoted $X \sim N(\mu, \sigma^2)$

**Linear transformation is also normally distributed:**

If $X \sim N(\mu, \sigma^2)$, then $a + bX \sim N(a + b\mu, b^2\sigma^2)$.

# 4.5 Specific probability distributions

Standardization of $X$ leads to standard normal distribution:

$$a = -\frac{\mu}{\sigma} \quad , \quad b = \frac{1}{\sigma}$$

$$z = \frac{x - \mu}{\sigma} \sim N(0, 1)$$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Thus, if $X \sim N(\mu, \sigma)$, then $f(x) = \frac{1}{\sigma}\Phi\left(\frac{x-\mu}{\sigma}\right)$.

# 4.5 Specific probability distributions

## The $\chi^2$ distribution:

$X$ is said to be $\chi^2(n)$ with $n$ degrees of freedom if

$$f_X(x) = \begin{cases} \frac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

If $z \sim N(0,1)$, then $x = z^2 \sim \chi^2(1)$.

If $z_i$ are **iid** $N(0,1)$ , then $\sum_{i=1}^{n} z_i^2 \sim \chi^2(n)$.