# 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, Chapter 6

# Online References

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

- Discrete RVs II: Functions of RV, conditional probabilities, specific distribution, total expectation theorem, joint probabilities
  https://www.youtube.com/watch?v=-qCEoqpwjf4
- Discrete RVs III: Conditional distributions and joint distributions continued
  https://www.youtube.com/watch?v=EObHWIEKGjA
- Multiple Continuous RVs: conditional pdf and cdf, joint pdf and cdf
  https://www.youtube.com/watch?v=CadZXGNauY0

# 4.6 Joint distributions

**Definition: Joint density function**

The joint density for two discrete random variables $X_1$ and $X_2$ is given as

$$f_{\boldsymbol{X}}(x_1, x_2) = \begin{cases} P(X_1 = x_{1i} \cap X_2 = x_{2i}) & \forall i, j \\ 0 & \text{else} \end{cases}$$

Properties:

- $1 \geq f_{\boldsymbol{X}}(x_1, x_2) \geq 0 \quad \forall \quad (x_1, x_2) \in \mathbb{R}^2$
- $\sum_{x_i} \sum_{x_j} f_X(x_{1i}, x_{2j}) = 1$

# 4.6 Joint distributions

ex: $X_1$: 1 A–M, 2 N–Z (lastname)

$X_2$: 1 Jan–June, 2 July–Dec (birthday)

$f_{X_1X_2}(x_1,x_2)$

joint distribution

marginal distribution

| $X_1$ \ $X_2$ | 1 | 2 | $f_{X_1}(x_1)$ |
|---|---|---|---|
| 1 | 0.2 | 0.25 | 0.45 |
| 2 | 0.3 | 0.25 | 0.55 |
| $f_{X_2}(x_2)$ | 0.5 | 0.5 | 1 |

$$P(X_1 \le 1, X_2 \le 2) = F(1,2) = 0.2 + 0.25 = 0.45$$

# 4.6 Joint distributions

**Definition: Joint cumulative distribution function**

The cdf for two discrete random variables $X_1$ and $X_2$ is given as

$$F_{\mathbf{X}}(x_1, x_2) = P(X_1 \leq x_1 \cap X_2 \leq x_2) = \sum_{x_{1i} \leq x_1} \sum_{x_{2i} \leq x_2} f_{\mathbf{X}}(x_{1i}, x_{2i})$$

it follows that

$$P(a \leq X_1 \leq b \cap c \leq X_2 \leq d) = \sum_{a \leq x_1 \leq b} \sum_{c \leq x_2 \leq d} f_{\mathbf{X}}(x_{1i}, x_{2i})$$

# 4.6 Joint distributions

If $X_1$ and $X_2$ are two continuous random variables, the following holds:

not a probability

$$\text{pdf} \qquad f_{\boldsymbol{X}}(x_1, x_2) = \frac{\partial^2 F_{\boldsymbol{X}}(x_1, x_2)}{\partial x_1 \partial x_2}$$

$$\text{cdf} \qquad F_{\boldsymbol{X}}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{\boldsymbol{X}}(u_1, u_2)\, du_2\, du_1$$

$$= P(X_1 \leq x_1 \cap X_2 \leq x_2)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, x_2)\, dx_1\, dx_2 = 1$$

$$f_{\boldsymbol{X}}(x_1, x_2) \geq 0$$

# 4.6 Joint distributions

$X_1, X_2$ are continuous r.v. p.d.f

$$f_X(x_1,x_2) = \frac{2}{5}(2x_1 + 3x_2) \quad \text{for} \quad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1$$

$$F_X(x_1,x_2) = \int_{-\infty \leftarrow 0}^{x_1} \int_{-\infty \leftarrow 0}^{x_2} \frac{2}{5}(2t_1 + 3t_2)\, dt_2\, dt_1$$

$$= \int_0^{x_1} \left[\frac{2}{5}\left(2t_1 t_2 + \frac{3}{2}t_2^2\right)\right]_0^{x_2} dt_1 = \int_0^{x_1} \frac{2}{5}\left(2t_1 x_2 + \frac{3}{2}x_2^2\right) dt_1$$

$$= \left[\frac{2}{5}\left(\frac{2x_2 t_1^2}{2} + \frac{3}{2}x_2^2 \cdot t_1\right)\right]_0^{x_1} = \frac{2}{5}\left(x_1^2 x_2 + \frac{3}{2}x_1 x_2^2\right)$$

# 4.7 Marginal Distributions

Derive the distribution of the individual variable from the joint distribution function:

$\rightarrow$ sum or integrate out the other variable

sum over all potential outcomes of $x_2$

$$f_{X_1}(x_1) = \begin{cases} \sum_{x_{2j}} f_{\boldsymbol{X}}(x_{1i}, x_{2j}) & \text{if } \boldsymbol{X} \text{ is discrete} \\ \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, x_2)\, dx_2 & \text{if } \boldsymbol{X} \text{ is continuous} \end{cases}$$

## 4.7 Marginal Distributions

ex: continuous

$$f_X(x_1,x_2) = \frac{2}{5}(2x_1+3x_2) \qquad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_X(x_1,x_2)\, dx_2 = \int_0^1 \frac{2}{5}(2x_1+3x_2)\, dx_2$$

$$= \left[\frac{2}{5}\left(2x_1x_2+\frac{3}{2}x_2^2\right)\right]_0^1 = \frac{2}{5}\left(2x_1+\frac{3}{2}\right) - 0$$

$$= \frac{4}{5}x_1 + \frac{3}{5} = f_{X_1}(x_1)$$

$$f_{X_2}(x_2) = \frac{2}{5} + \frac{6}{5}x_2$$

# 4.7 Marginal Distributions

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_X(x_1, x_2)\,dx_1 = \int_0^1 \frac{2}{5}(2x_1 + 3x_2)\,dx_1$$

$$= \left[\frac{2}{5}\left(\cancel{\tfrac{2}{2}}\,x_1^2 + 3x_1x_2\right)\right]_0^1 = \frac{2}{5}\left(1 + 3x_2\right) = \frac{6}{5}x_2 + \frac{2}{5}$$

# 4.7 Stochastic Independence

Two random variables are stochastically independent if their joint density is the product of the marginal densities:

not

$$f_{\mathbf{X}}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \Leftrightarrow X_1 \text{ and } X_2 \text{ are independent.}$$

stochastically independent

Under independence the cdf factors as well:

$$F_{\mathbf{X}}(x_1, x_2) = F_{X_1}(x_1) \cdot F_{X_2}(x_2).$$

Expectations in a joint distribution are computed with respect to the marginals.

| | 1 | 2 | |
|---|---|---|---|
| 1 | 0.2 | 0.25 | $\boxed{0.45}$ |
| 2 | 0.3 | 0.25 | 0.55 |
| | $\boxed{0.5}$ | 0.5 | 1 |

| | 1 | 2 | |
|---|---|---|---|
| 1 | 0.225 | 0.225 | 0.45 |
| 2 | 0.275 | 0.275 | 0.55 |
| | 0.5 | 0.5 | 1 |

$0.5 \cdot 0.45 = 0.225 \neq 0.2$

# 4.7 Stochastistic Independence

Two random variables are stochatistically independent if their joint density is the product of the marginal densities:

$$f_{\boldsymbol{X}}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \Leftrightarrow X_1 \text{ and } X_2 \text{ are independent.}$$

Under independence the cdf factors as well:

$$F_{\boldsymbol{X}}(x_1, x_2) = F_{X_1}(x_1) \cdot F_{X_2}(x_2).$$

Expectations in a joint distribution are computed with respect to the marginals.

# 4.7 Stochatistic Independence

ex: continuous r.v.

$f_{\mathbf{x}}(x_1, x_2) = f_{x_1}(x_1) \cdot f_{x_2}(x_2)$ check this

$f_{\mathbf{x}}(x_1, x_2) = \frac{3}{2} x_1^2 \cdot x_2$ for $0 \le x_1 \le 1$, $0 \le x_2 \le 2$

1. get marginal densities $f_{x_1}(x_1)$ and $f_{x_2}(x_2)$

$$f_{x_1}(x_1) = \int_0^2 f_{x_1 x_2}(x_1, x_2)\, dx_2 = \int_0^2 \frac{3}{2} x_1^2 x_2 \, dx_2 = \frac{3}{2} x_1^2 \left[\frac{1}{2} x_2^2\right]_0^2$$

$$= \frac{3}{4} x_1^2 \cdot 4 - 0 = 3 x_1^2$$

$$f_{x_2}(x_2) = \int_0^1 f_{x_1 x_2}(x_1, x_2)\, dx_1 = \int_0^1 \frac{3}{2} x_1^2 x_2 \, dx_1 = \frac{3}{2} x_2 \left[\frac{1}{3} x_1^3\right]_0^1$$

$$= \frac{1}{2} x_2 - 0 = \frac{1}{2} x_2$$

→ YES, INDEPENDENT

check: $f_{x_1}(x_1) \cdot f_{x_2}(x_2) = 3 \cdot x_1^2 \cdot \frac{1}{2} x_2 = \frac{3}{2} x_1^2 x_2 = f_{\mathbf{x}}(x_1, x_2)$

# 4.8 Random Samples

$\{X_1, \dots, X_n\}$ if $X_i$ is iid

independent identically distributed

joint density

independence

$$f_{X_1 X_2 \dots X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \dots \cdot f_{X_n}(x_n)$$

$$= f_X(x_1) \cdot f_X(x_2) \cdot \dots \cdot f_X(x_n)$$

$$= \prod_{i=1}^{n} f_X(x_i)$$

same marginal distribution for all

# 4.9 Covariance and correlation

$$Cov[X_1, X_2] = E[(X_1 - E[X_1])(X_2 - E[X_2])]$$

Properties:

- symmetry: $Cov[X_1, X_2] = Cov[X_2, X_1]$
- linear transformation:

$$Y_1 = b_0 + b_1 X_1 \qquad Y_2 = c_0 + c_1 X_2$$
$$\Rightarrow Cov[Y_1, Y_2] = b_1 c_1 Cov[X_1, X_2]$$

- calculation:

$$Cov[X_1, X_2] = \begin{cases} \sum_{x_{1i}} \sum_{x_{2j}} x_{1i} x_{2j} f_{\mathbf{X}}(x_{1i}, x_{2j}) - E[X_1]E[X_2] \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{X}}(x_1, x_2)\, dx_2\, dx_1 - E[X_1]E[X_2] \end{cases}$$

# 4.9 Covariance and correlation

$$\mathrm{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))] = E(XY) - E(X)\cdot E(Y)$$

multiply out

$$= E[XY - E(X)Y - XE(Y) + E(X)E(Y)]$$

seperate terms

$$= E[XY] - E[E(X)Y] - E[XE(Y)] + E[E(X)E(Y)]$$

constants outside

$$= E[XY] - E[X]E(Y) - E(\cancel{X})E(Y) + E(\cancel{X})E(Y)$$

$$= E[XY) - E(X)E(Y)$$

$\rightarrow$ shown with $E(\cdot)$ operator notation

# 4.9 Covariance and correlation

continuous r.v. X and Y with integrals:

$$\mathrm{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-E(X))(y-E(Y))\, f_{XY}(x,y)\, dx\, dy$$

multiply out

split up integral

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy - E(X)y - xE(Y) + xy)\, f_{XY}(x,y)\, dx\, dy$$

constants outside

$$= \underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, f_{XY}(x,y)\, dx\, dy}_{E(XY)} - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(X)\, y\, f_{XY}(x,y)\, dx\, dy$$

$$- \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, E(Y)\, f_{XY}(x,y)\, dx\, dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(X)E(Y)\, f_{XY}(x,y)\, dx\, dy$$

familiar terms

$$= E(XY) - E(X)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{XY}(x,y)\, dx\, dy - E(Y)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{XY}(x,y)\, dx\, dy$$

constant w.r.t. x-integral

constant w.r.t. y-integral

$$+ E(X)E(Y)\underbrace{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\, dy}_{=1}$$

# 4.9 Covariance and correlation

$$= E(XY) - E(X)\int_{-\infty}^{\infty} y \underbrace{\left[\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\right]}_{=f_Y(y)} dy - E(Y)\int_{-\infty}^{\infty} x \underbrace{\left[\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\right]}_{=f_X(x)} dx + E(X)E(Y)$$

integrate out y: marginal dist. of X

$$= E(XY) - E(X)\underbrace{\int_{-\infty}^{\infty} y\, f_Y(y)\,dy}_{=E(Y)} - E(Y)\underbrace{\int_{-\infty}^{\infty} x\, f_X(x)\,dx}_{=E(X)} + E(X)E(Y)$$

$$= E(XY) - E(X)E(Y)$$

## 4.9 Covariance and correlation

If $X$ and $Y$ are independent, then $Cov(X,Y) = 0$

$$\underbrace{f_{XY}(x,y) = f_X(x) \cdot f_Y(y)}$$ $\rightarrow$ $E(XY) = E(X)E(Y)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \, \underbrace{f_{XY}(x,y)}_{= f_X(x) \cdot f_Y(y)} \, dy \, dx$$

independence

$$= \int_{-\infty}^{\infty} x f_X(x) dx \cdot \int_{-\infty}^{\infty} y f_Y(y) dy$$

$$= E(X) \cdot E(Y)$$

$$Cov(X,Y) = \underbrace{E(XY)}_{=E(X) \cdot E(Y)} - E(X) \cdot E(Y)$$

independence

$$= 0$$

# 4.9 Covariance and correlation

**Pearson's correlation coefficient**

$$\rho_{X_1,X_2} = \frac{Cov(X_1,X_2)}{\sqrt{Var(X_1)\cdot Var(X_2)}} = \frac{\sigma_{X_1,X_2}}{\sigma_{X_1}\sigma_{X_2}} \quad \in [-1,1]$$

- If $X_1$ and $X_2$ are independent, they are also uncorrelated.
- Uncorrelated does not imply independence!
- Exception: normal distribution, characterized by 1st and 2nd moment.

$$\boxed{f_X(x_1,x_2) = f_{X_1}(x_1)\cdot f_{X_2}(x_2)} \Longrightarrow \quad \rho_{X,Y} = 0$$

$X_1$ and $X_2$ are independent $\quad \not\Longleftarrow$

# 4.10 Conditional Distributions

- Distribution of the variable $X_1$ given that $X_2$ takes on a certain value $x_1$.
- Closely related to conditional probabilities:

$$P(X_1 = x_1 | \boxed{X_2 = x_2}) = \frac{P(X_1 = x_1 \cap X_2 = x_2)}{\boxed{P(X_2 = x_2)}}$$

conditional pdf of $X_1$ given $X_2 = x_2$:

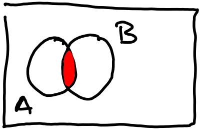
$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1, x_2)}{\boxed{f_{X_2}(x_2)}}$$

If $X_1$ and $X_2$ are independent: $f_{X_1X_2}(x_1,x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$

$$f_{X_1|X_2}(x_1,x_2) = \frac{f_{X_1}(x_1)\cdot \cancel{f_{X_2}(x_2)}}{\cancel{f_{X_2}(x_2)}} = f_{X_1}(x_1)$$

# 4.10 Conditional Distributions

conditional cdf of $X_1$ given $X_2 = x_2$:




$$P(\underline{X_1 = x_1}|\boxed{X_2 = x_2}) = \sum_{x_{1i} \leq x_1} f_{X_1|X_2}(x_{1i}|x_2) = F_{X_1|X_2}(x_1|x_2)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

If $X_1$ and $X_2$ are independent, the conditional probability and the marginal probability coincide:

$$f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$$

univariate function

because

$$f_{X_1 X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

$X_2 = 1$

| $X_1$ \ $X_2$ | 1 (Jan-Jun) | 2 (July-Dec) | |
|---|---|---|---|
| → A-M 1 | 0.2 | 0.25 | 0.45 |
| → N-Z 2 | 0.3 | 0.25 | 0.55 |
| | 0.5 | 0.5 | 1 |

$$f(X_1 = 1 | X_2 = 1) = \frac{0.2}{0.5} = \frac{f(X_1 = 1, X_2 = 1)}{f(X_2 = 1)}$$

$$f(X_1 = 2 | X_2 = 1) = \frac{f(X_1 = 2, X_2 = 1)}{f(X_2 = 1)} = \frac{0.3}{0.5}$$

# 4.10 Conditional Distributions

ex: $f_{X_1X_2}(x_1,x_2) = \frac{2}{5}(2x_1 + 3x_2)$ for $0 \le x_1 \le 1$, $\boxed{0 \le x_2 \le 1}$

$x_1 = 0$

consider

$$f_{\boxed{X_2}|\boxed{X_1=0}}(x_2|x_1=0) = \frac{f_{X_1X_2}(x_1=0,\ x_2)}{f_{X_1}(x_1=0)}$$

not a probability

marginal density from before:

$f_{X_1}(x_1) = \frac{4}{5}x_1 + \frac{3}{5}$ for $x_1 = 0$ $f_{X_1}(x_1=0) = \frac{3}{5}$

$f_{X_1X_2}(0,x_2) = \frac{6}{5}x_2$

$$f_{X_2|X_1=0}(x_2|x_1=0) = \frac{\frac{6}{5}x_2}{\frac{3}{5}} = 2x_2$$

function of $x_2$: univariate function only depending on $x_2$

# 4.10 Conditional Distributions

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1X_2}(x_1,x_2)}{f_{X_2}(x_2)}$$

The joint pdf can be derived from conditional and marginal densities in 2 ways:

ATS

$$f_{X_1X_2} = f_{X_1|X_2}(x_1|x_2) \cdot f_{X_2}(x_2) = f_{X_2|X_1}(x_2|x_1) \cdot f_{X_1}(x_1)$$

$$\Leftrightarrow f_{X_1X_2}(x_1,x_2) = f_{X_1|X_2}(x_1|x_2) \cdot f_{X_2}(x_2)$$

# 4.11 Conditional Moments

function of just y

$$E[Y|X] = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dy$$

e.g. $g(Y) = Y^2$

$$E[Y^k|X=x] = \sum_j \boxed{y_j^k} \cdot \boxed{\frac{P(X = x \cap Y = y_j)}{P(X=x)}}$$

$$= \sum_{y_j} \boxed{y_j^k} \cdot P(Y = y_j | X = x)$$

$$= \sum_{y_j} \boxed{y_j^k} \cdot f_{Y|X}(y_j|x)$$

$$= \sum_{y_j} y_j^k \cdot \frac{f_{XY}(x, y_j)}{f_X(x)} \qquad \text{if } Y \text{ is discrete}$$

for conditional moments: weights are altered to conditional densities

$f_{Y|X=x}(X=x)$

$$E[Y^k|X=x] = \underbrace{\int_{-\infty}^{\infty} y^k \cdot \frac{f_{XY}(x, y)}{f_X(x)}\, dy}_{} \qquad \text{if } Y \text{ is continuous}$$

# 4.11 Conditional Moments

$$Var[Y|X=x] = E_{Y|X}[(Y - E[Y|X=x])^2]$$

$$= \sum_{y_j} (y_j - E[Y|X=x])^2 \cdot f_{Y|X}(y_j|x)$$

if $Y$ is discrete

$$Var[Y|X=x] = E_{Y|X}[(Y - E[Y|X=x])^2]$$

$$= \int_{-\infty}^{\infty} (y - E[Y|X=x])^2 \cdot f_{Y|X}(y|x)\,dy$$

if $Y$ is continuous

# 4.11 Conditional Moments

## Law of Total Expectations/ Law of Iterated Expectations

constant, not random variable

$X=x$ ex: $X=0$

$E_X\left[E_{Y|X}(Y|\cancel{X=x})\right]$

if $X=x$ is still a r.v.

$E_{Y|X}(Y|X=x)$ is still a r.v. w.r.t. $X$

$$E[Y] = E_X\left[E[Y|X]\right]$$

$$E_X\left[E_{Y|X}[Y|X]\right] = E[Y] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y \cdot \frac{f_{XY}(x,y)}{f_X(x)} dy\right] f_X(x)\, dx$$

$E_{Y|X}$ is a random value as $X$ is a random variable.

# 4.11 Conditional Moments

## Law of Total Expectations/ Law of Iterated Expectations

LTE - Law of Total Expectations

or LIE - Law of Iterated Expectations

so far: $E(x)$ is constant so $E_x[E(x)] = E(x)$

$$E_x[\underbrace{E_{Y|X}(Y|X)}_{\text{why is this a r.v.?}}] = E_Y(Y)$$

$E_{Y|X}(Y|X=x)$

→ if $x$ is already realized, this is a constant

→ if $X$ is <u>not</u> realized yet, this is a r.v.

# 4.11 Conditional Moments

## Law of Total Expectations/ Law of Iterated Expectations

$$E_X\left[E_{Y|X}(Y|X)\right] \quad \text{for continuous r.v.}$$

$$= \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} y\, f_{Y|X}(y|x) \cdot dy}_{\text{r.v. w.r.t. } X = g(x)} \cdot f_X(x) \cdot dx$$

change order of integration

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, \frac{f_{XY}(x,y)}{\cancel{f_X(x)}}\, dy\, \cancel{f_X(x)}\, dx$$

$$= \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f_{XY}(x,y)\, dx}_{=f_Y(y) \text{ marginal distribution}} dy = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy = E(Y)$$

# 4.11 Conditional Moments

## Law of Total Expectations/ Law of Iterated Expectations

## 4.12 The bivariate normal distribution

**Definition: Bivariate normal distribution**

Two random variables $X_1$ and $X_2$ are jointly normally distributed if they are described by the joint pdf

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left[-\frac{1}{2}q(x_1, x_2)\right]$$

where

$$q(x_1, x_2) = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right].$$

Key: the normal distribution always reproduces itself

$X_1, \dots, X_n$ jointly normally distribution

# 4.12 The bivariate normal distribution

If $(X_1, X_2) \sim \boxed{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then joint

- $f_{X_1}(x_1) \sim \boxed{N}(\mu_1, \sigma_1^2)$, $X_1$ marginal
  $f_{X_2}(x_2) \sim \boxed{N}(\mu_2, \sigma_2^2)$, $X_2$
- $f_{X_1|X_2} \sim \boxed{N}(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1 - \rho^2))$, conditional
  $f_{X_2|X_1} \sim \boxed{N}(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1 - \rho^2))$.

$Z = aX_1 + bX_2 \sim \boxed{N}\left(E(aX_1 + bX_2), Var(aX_1 + bX_2)\right)$

# 4.12 Linear combination of normal distributed r.v.s

If $X$ and $Y$ are normally distributed, then

$$Z = aX + bY \sim N(E(Z), Var(Z))$$

linear combination

with

$$E(Z) = E[aX + bY] = a \cdot E(X) + b \cdot E(Y) \quad (*)$$

split up sum

constants outside of $E(\cdot)$

$$Var(Z) = Var(aX+bY) = E[(Z - E(Z))^2]$$

$$= E[((aX+bY) - E(aX+bY))^2]$$

$$\overset{(*)}{=} E[(aX+bY - (aE(X)+bE(Y)))^2]$$

$$= E[(\underbrace{a(X-E(X))}_{A} + \underbrace{b(Y-E(Y))}_{B})^2]$$

$$(A+B)^2 = A^2 + 2AB + B^2$$

multiply out

$$= E\left[a^2 (X-E(X))^2 + 2ab(X-E(X))(Y-E(Y)) + b^2 (Y-E(Y))^2\right]$$

split up sum

$$= E\left[a^2 (X-E(X))^2\right] + E\left[2ab(X-E(X))(Y-E(Y))\right] + E\left[b^2 (Y-E(Y))^2\right]$$

constants outside

familiar terms

$$= a^2 \underbrace{E\left[(X-E(X))^2\right]}_{=Var(X)} + 2ab \underbrace{E\left[(X-E(X))(Y-E(Y))\right]}_{Cov(X,Y)} + b^2 \underbrace{E\left[(Y-E(Y))^2\right]}_{Var(Y)}$$

$$= a^2 Var(X) + b^2 Var(Y) + \underbrace{2ab\, Cov(X,Y)}$$

$$= Var(aX+bY)$$

basis: portfolio diversification

$$Z = aX+bY \sim \boxed{N}(E(Z), Var(Z))$$

$$X \sim \boxed{N}(E(X), Var(X))$$

$$Y \sim \boxed{N}(E(Y), Var(Y))$$

binomial formula: $(A-B)^2 = A^2+B^2-2AB$

note: $W = aX - bY$

$$Var(W) = a^2 Var(X) + b^2 Var(Y) - 2ab\, Cov(X,Y)$$

# 4.13 Multivariate Distributions

$\boldsymbol{x}$ a random vector with joint density $f_{\boldsymbol{X}}(\boldsymbol{x})$

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f_{\boldsymbol{X}}(\boldsymbol{t})\, dt_1\, dt_2 \ldots dt_{n-1}\, dt_n$$

Expected Value:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{pmatrix}$$

# 4.13 Multivariate Distributions

## Covariance Matrix

$$E\left[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})'\right]$$

$$= \begin{pmatrix} (x_1-\mu_1)(x_1-\mu_1) & (x_1-\mu_1)(x_2-\mu_2) & \dots & (x_1-\mu_1)(x_n-\mu_n) \\ (x_2-\mu_2)(x_1-\mu_1) & (x_2-\mu_2)(x_2-\mu_2) & \dots & (x_2-\mu_2)(x_n-\mu_n) \\ \vdots & & & \\ (x_n-\mu_n)(x_1-\mu_1) & (x_n-\mu_2)(x_n-\mu_2) & \dots & (x_n-\mu_n)(x_n-\mu_n) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots \sigma_{2n} \\ \vdots & & \\ \sigma_{n1} & \dots & \dots \sigma_n^2 \end{pmatrix} = E\left[\boldsymbol{x}\boldsymbol{x}'\right] - \boldsymbol{\mu}\boldsymbol{\mu}' = \boldsymbol{\Sigma}$$

# 4.13 Multivariate Distributions

## Linear Transformation: sum of $n$ random variables $\sum_{i=1}^{n} a_i x_i$

$$\begin{aligned}
E[a_1 x_1 + a_2 x_2 + \dots a_n x_n] &= E[\boldsymbol{a}'\boldsymbol{x}] \\
&= \boldsymbol{a}' E[\boldsymbol{x}] = \boldsymbol{a}'\boldsymbol{\mu} \\
Var[\boldsymbol{a}'\boldsymbol{x}] &= E[(\boldsymbol{a}'\boldsymbol{x} - E[\boldsymbol{a}'\boldsymbol{x}])^2] \\
&= E[(\boldsymbol{a}'(\boldsymbol{x} - E[\boldsymbol{x}])^2] \\
&= E[(\boldsymbol{a}'(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{a}] \\
&= \boldsymbol{a}' E[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})']\boldsymbol{a} \\
&= \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \sigma_{ij}
\end{aligned}$$

# 4.13 Multivariate Distributions

## Linear transformation: $\boldsymbol{y} = \boldsymbol{Ax}$

$i$-th element in $\boldsymbol{y} = \boldsymbol{Ax}$ is $y_i = \boldsymbol{a_i x}$ with $\boldsymbol{a_i}$ $i$-th row in $\boldsymbol{A}$

$\Rightarrow E[y_i] = E[\boldsymbol{a_i x}] = \boldsymbol{a_i \mu}$ as before

$$
\begin{aligned}
E[\boldsymbol{y}] &= E[\boldsymbol{Ax}] = \boldsymbol{A}E[\boldsymbol{x}] = \boldsymbol{A\mu} \\
Var[\boldsymbol{y}] &= E[(\boldsymbol{y} - E[\boldsymbol{y}])(\boldsymbol{y} - E[\boldsymbol{y}])'] \\
&= E[(\boldsymbol{Ax} - \boldsymbol{A\mu})(\boldsymbol{Ax} - \boldsymbol{A\mu})'] \\
&= E[(\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{\mu})[(\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{\mu})]'] \\
&= E[\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{A}'] \\
&= \boldsymbol{A}E[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})']\boldsymbol{A}' = \boldsymbol{A\Sigma A}'
\end{aligned}
$$