# Probability and Statistics

## 3 – Discrete Random Variables

Stefan Heiss

Technische Hochschule Ostwestfalen-Lippe
Dep. of Electrical Engineering and Computer Science

November 3, 2023

# Independent Random Variables

## Definition (3.26)

Let $X, Y$ be discrete random variables with respect to the same probability measure Pr, i.e. with respect to the same triple $(\Omega, \mathcal{A}, \Pr)$. Then $X$ and $Y$ are called *independent random variables* if for all subsets $S, T \subseteq \mathbb{R}$ the events $X^{-1}(S)$ and $Y^{-1}(T)$ are independent, i.e.:

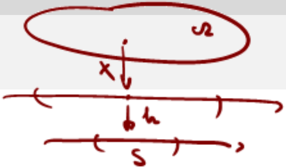
$$\Pr(X^{-1}(S) \cap Y^{-1}(T)) = \Pr(X \in S, Y \in T) = \Pr(X \in S) \cdot \Pr(Y \in T) = \Pr(X^{-1}(S)) \cdot \Pr(Y^{-1}(T))$$

More generally, any finite number of random variables $X_1, \ldots, X_n$ are *independent*, if

$$\Pr\left(\bigcap_{i=1}^{n} \underbrace{\{X_i \in S_i\}}_{X_i^{-1}(S_i)}\right) = \prod_{i=1}^{n} \Pr(X_i \in S_i)$$

for all subsets $S_1, \ldots, S_n \subseteq \mathbb{R}$.

# Independent Random Variables




**Lemma (3.27)**

*Let $X$ and $Y$ be independent random variables and $h, k : \mathbb{R} \to \mathbb{R}$. Then also $h(X)$ and $k(Y)$ are independent random variables.*

Let $S, T \subseteq \mathbb{R}$

$$\Pr(h(X) \in S, k(Y) \in T) = \Pr\left((h \circ X)^{-1}(S) \cap (k \circ Y)^{-1}(T)\right)$$

$$= \Pr\left(X^{-1}(\underline{h^{-1}(S)}) \cap Y^{-1}(\underline{k^{-1}(T)})\right) = \Pr\left(X^{-1}(h^{-1}(S))\right) \cdot \Pr\left(Y^{-1}(k^{-1}(T))\right)$$

$$= \Pr\left((h(X))^{-1}(S)\right) \cdot \Pr\left((k(Y))^{-1}(T)\right) = \Pr(h(X) \in S) \cdot \Pr(k(Y) \in T)$$

$\square$

# Independent Random Variables

Lemma (3.28)

*Let $X, Y$ be discrete random variables with respect to the same probability measure* Pr, *i.e. with respect to the same triple* $(\Omega, \mathcal{A}, \mathrm{Pr})$. *Then the following conditions are equivalent:*

(i) $X$ *and* $Y$ *are independent* $\Pr(X \in S, Y \in T) = \Pr(X \in S) \cdot \Pr(Y \in T)$ f.a. $S, T \subseteq \mathbb{R}$

$\Pr(X=x, Y=y) = \Pr(X \in \{x\}, Y \in \{y\}) = \Pr(X \in \{x\}) \cdot \Pr(Y \in \{y\}) = p_X(x) \cdot p_Y(y)$

(ii) $p_{XY}(x, y) = p_X(x) \cdot p_Y(y)$ *for all* $x, y \in \mathbb{R}$

(iii) $E(h(X) \cdot k(Y)) = E(h(X)) \cdot E(k(Y))$ *for all functions* $h, k : \mathbb{R} \to \mathbb{R}$

# Proof of (3.28)

In order to prove the lemma, it suffices to show:

$$\text{(i)} \implies \text{(ii)} \implies \text{(iii)} \implies \text{(i)}$$

(i) $\implies$ (ii): Trivial, as (ii) simply states the independence of events $X^{-1}(S)$ and $Y^{-1}(T)$, where $S$ and $T$ just consist of single elements: $S = \{x\}$, $T = \{y\}$.

see last slide !

# Proof of (3.28)

(ii) $\Longrightarrow$ (iii):

$$
\begin{aligned}
E(h(X)\cdot k(Y)) &= \sum_{(x,y)\in\mathbb{R}^2} \underline{h(x)\,k(y)} \cdot p_{XY}(x,y) \\
&= \sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}} h(x)\,k(y)\cdot p_X(x)\,p_Y(y) \qquad \text{(by (ii))} \\
&= \left(\sum_{x\in\mathbb{R}} h(x)\,p_X(x)\right)\left(\sum_{y\in\mathbb{R}} k(y)\,p_Y(y)\right) \\
&= E(h(X))\cdot E(k(Y))
\end{aligned}
$$

by (3.24)(ii) with

$g:\mathbb{R}^2\to\mathbb{R}$, $g:(x,y)\mapsto h(x)\cdot k(y)$

# Proof of (3.28)

(iii) $\implies$ (i): Let $S, T \subseteq \mathbb{R}$. Consider the *indicator functions* $I_S : \mathbb{R} \to \{0,1\}$ and $I_T : \mathbb{R} \to \{0,1\}$ defined by:

$$I_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases} \qquad \text{and} \qquad I_T(x) = \begin{cases} 1 & \text{if } x \in T \\ 0 & \text{if } x \notin T \end{cases}$$

Then:

$$\begin{aligned} \Pr(X \in S) \cdot \Pr(Y \in T) &= \left(\sum_{x \in S} p_X(x)\right)\left(\sum_{y \in T} p_Y(y)\right) \\ &= \left(\sum_{x \in \mathbb{R}} I_S(x) p_X(x)\right)\left(\sum_{y \in \mathbb{R}} I_T(y) p_Y(y)\right) \end{aligned}$$

# Proof of (3.28)

$$
\begin{aligned}
\left(\sum_{x\in\mathbb{R}} I_S(x)p_X(x)\right)\left(\sum_{y\in\mathbb{R}} I_T(y)p_Y(y)\right) &= E(I_S(X))\cdot E(I_T(Y)) \\
&= E(I_S(X)\cdot I_T(Y)) && \text{(by (iii))} \\
&= \sum_{(x,y)\in\mathbb{R}^2} I_S(x)\cdot I_T(y)\cdot p_{XY}(x,y) && \text{(3.24) (ii)} \\
&= \sum_{x\in S}\sum_{y\in T} p_{XY}(x,y) \\
&= \Pr(X\in S, Y\in T)
\end{aligned}
$$

# Independent Random Variables

**Corollary (3.29)**

(i) *Let $X, Y$ be independent random variables. Then:*

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

(3.28)(iii) with

$h = k = \mathrm{Id}_{\mathbb{R}}$

(ii) *Let $X, Y$ be independent random variables with moment generating functions $\phi_X(t)$ and $\phi_Y(t)$. Then:*

$$\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t)$$

$h = k : x \mapsto e^{x \cdot t}$

(iii) *Let $X_1, \ldots, X_n$ be independent random variables having all the same distribution. Then:*

$$\phi_{X_1+\cdots+X_n}(t) = (\phi_{X_1}(t))^n$$

# Correlation

**Definition (3.30)**

The *correlation* between two random variables $X$ and $Y$ is defined to be $E(XY)$.

# Cauchy–Schwarz Inequality

Lemma (3.31)

*Let $X, Y$ be random variables. Then:*

$$\boxed{|E(XY)| \leq \sqrt{E(X^2) \cdot E(Y^2)}}$$

*If $\mathcal{A} = \mathcal{P}(\Omega)$, then equality holds if and only if there exists some $\lambda \in \mathbb{R}$ with:*

$$X(\omega) = \lambda \cdot Y(\omega) \qquad \text{for all } \underline{\omega \in \Omega^*} := \underline{\{\omega \in \Omega \mid \Pr(\{\omega\}) > 0\}}$$

$\left(\textit{In this case: } \lambda = \frac{E(XY)}{E(Y^2)}\right)$

# Proof of (3.31)

Put $\lambda := \frac{E(XY)}{E(Y^2)}$. Then:

$$
\begin{array}{rrcl}
 & 0 & \leq & E((X-\lambda Y)^2) \;=\; E(X^2 - 2\lambda XY + \lambda^2 Y^2) \\
\Longrightarrow & 0 & \leq & E(X^2) \;-\; 2\lambda\, E(XY) \;+\; \lambda^2\, E(Y^2) \\
\Longrightarrow & 0 & \leq & E(X^2) \;-\; 2\frac{(E(XY))^2}{E(Y^2)} \;+\; \frac{(E(XY))^2}{(E(Y^2))^{\cancel{2}}} \cancel{E(Y^2)} \\
\Longrightarrow & 0 & \leq & E(X^2) \;-\; \frac{(E(XY))^2}{E(Y^2)} \\
\Longrightarrow & (E(XY))^2 & \leq & E(X^2)\, E(Y^2) \\
\Longrightarrow & |E(XY)| & \leq & \sqrt{E(X^2)\cdot E(Y^2)}
\end{array}
$$

$|E(XY)| = \sqrt{E(X^2)\cdot E(Y^2)}$

$\Rightarrow 0 = E((X-\lambda Y)^2)$

$\Rightarrow (X-\lambda Y)|_{\Omega'} = 0 \Rightarrow X|_{\Omega'} = \lambda Y|_{\Omega'}$

# Covariance

## Definition (3.32)

Let $X, Y$ be random variables and $m_X = E(X)$, ${\sigma_X}^2 = Var(X)$, $m_Y = E(Y)$, ${\sigma_Y}^2 = Var(Y)$.
The *covariance* between $X$ and $Y$ is defined by:

$= E(XY - m_Y X - m_X \cdot Y + m_X m_Y)$

$$Cov(X, Y) \;:=\; E\left((X - m_X)(Y - m_Y)\right) \;=\; \underline{E(XY)} - \underline{m_X\, m_Y}$$

$X$ and $Y$ are said to be *uncorrelated* if $Cov(X, Y) = 0$.

# $X, Y$ independent $\implies$ $X, Y$ uncorrelated

**Remark (3.33)**

Let $X, Y$ be random variables. If $X$ and $Y$ are independent, then $X$ and $Y$ are also uncorrelated.

Pf.: See (3.23) (i)

# Normalized Random Variable

**Definition (3.34)**

Let $X$ be a random variable whose expectation $m = E(X)$ and standard deviation $\sigma_X$ exist and are finite. Then, the *normalized random variable* defined by $X$ is:

$$X^o = \frac{X - m}{\sigma_X}$$

Note: $E(X^o) = 0$ and $\sigma_{X^o} = 1$. ← Exc.: Check it!

# Correlation Coefficient

**Definition (3.35)**

The *correlation coefficient* of $X$ and $Y$ is defined to be the correlation between the normalized random variables defined by $X$ and $Y$:

$$\rho_{XY} \;:=\; E\left(\left(\frac{X - m_X}{\sigma_X}\right)\left(\frac{Y - m_Y}{\sigma_Y}\right)\right) \;=\; \frac{Cov(X, Y)}{\sigma_X\,\sigma_Y}$$

$$Cov(X,Y) = E((X - m_X)(Y - m_Y))$$

# Correlation Coefficient

Lemma (3.36)

*Let* $X, Y$ *be random variables. Then:*

$$\boxed{|\rho_{XY}| \quad \leq \quad 1}$$

*If* $\mathcal{A} = \mathcal{P}(\Omega)$, *then equality holds if and only if* $X|_{\Omega^*}$ *and* $Y|_{\Omega^*}$ *are related by a linear function plus a constant, where* $\Omega^* := \{\omega \in \Omega \mid \Pr(\{\omega\}) > 0\}$.

# Proof of (3.36)

$$
\begin{aligned}
|\rho_{XY}| &= \left| E\left( \left( \frac{X - m_X}{\sigma_X} \right) \left( \frac{Y - m_Y}{\sigma_Y} \right) \right) \right| \\
&\overset{(3.31)}{\leq} \sqrt{ E\left( \left( \frac{X - m_X}{\sigma_X} \right)^2 \right) E\left( \left( \frac{Y - m_Y}{\sigma_Y} \right)^2 \right) } = 1
\end{aligned}
$$

Furthermore, $|\rho_{XY}| = 1$, if and only if

$$
\frac{X|_{\Omega^*} - m_X}{\sigma_X} = \lambda \cdot \frac{Y|_{\Omega^*} - m_Y}{\sigma_Y} \qquad (\lambda = \rho_{XY} \in \{1, -1\})
$$

$$
\iff \qquad X|_{\Omega^*} = \rho_{XY} \cdot \frac{\sigma_X}{\sigma_Y} \cdot Y|_{\Omega^*} + \left( m_X - \rho_{XY} \cdot \frac{\sigma_X}{\sigma_Y} m_Y \right)
$$

# $Var(X_1 + \cdots + X_n)$

## Theorem (3.37)

*Let $X_1, \dots, X_n$ be random variables. Then:*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2 \cdot \sum_{1 \leq i < j \leq n} Cov(X_i, X_j)$$