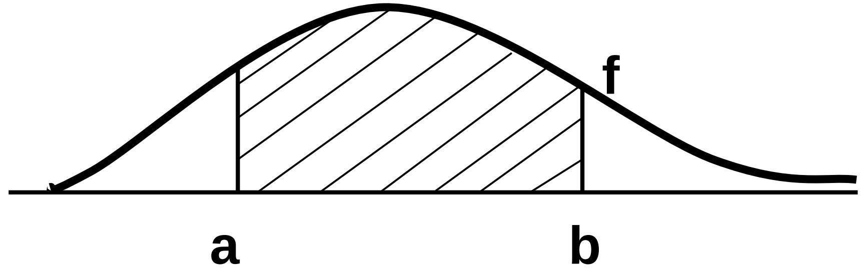
Slides Probability Theory
Chapter 7

Ch. 7: 2 or more continuous variables

- Independence
- The distribution of functions of 2 or more r.v.`s, such as $X+Y$ and $\max(X,Y)$
- The Central Limit Theorem and applications

DISCRETE	CONTINUOUS
S_X is numerable **probability** $P(X=x)$ $\sum_x P(X=x)=1$	 **"probability"** $x\cdot f(x)dx$ $\int_{-\infty}^{\infty} f(x)dx=1$
$P(a<X<b)=\sum_{x\in(a,b)} P(X=x)$	$P(a<X<b)=\int_a^b f(x)dx$
$E(X)=\sum_x x\cdot P(X=x)$	$E(X)=\int_{-\infty}^{\infty} x\cdot f(x)dx$
$E(X^2)=\sum_x x^2\cdot P(X=x)$	$E(X^2)=\int_{-\infty}^{\infty} x^2\cdot f(x)dx$

Independence of X and Y

General definition: **X and Y are independent** if for all subsets A and B of $\mathbb{R}$ the following equality holds:

$$\boldsymbol{P}(\boldsymbol{X} \in \boldsymbol{A} \text{ and } \boldsymbol{Y} \in \boldsymbol{B}) = \boldsymbol{P}(\boldsymbol{X} \in \boldsymbol{A}) \times \mathbf{P}(\boldsymbol{Y} \in \boldsymbol{B})$$

Discrete: choose $A = \{\boldsymbol{x}\}$ and $B = \{y\}$, then

$$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x} \text{ and } \boldsymbol{Y} = \boldsymbol{y}) = \boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}) \times \boldsymbol{P}(\boldsymbol{Y} = \boldsymbol{y})$$

Continuous: simply use the general definition,

Example: X and Y are independent and $Exp\left(\lambda = \frac{1}{2}\right)$

Then: $P(X > 2 \text{ and } Y < 4) = P(X > 2) \times P(Y < 4)$

$$= \quad e^{-\frac{1}{2}\cdot 2} \quad \times \left(1 - e^{-\frac{1}{2}\cdot 4}\right)$$

$$= e^{-1} - e^{-3} \approx 31.8\%$$

Example: maximum of 2 waiting times *X* and *Y*

- Two customers are served independently: whose service is completed first, $\boldsymbol{min}(\boldsymbol{X},\boldsymbol{Y})$ whose last, $\boldsymbol{max}(\boldsymbol{X},\boldsymbol{Y})$?
- **Model:** waiting times are independent and both $Exp(\lambda)$.
- $P(\mathbf{max}(\boldsymbol{X},\boldsymbol{Y}) \leq x) = P(X \leq x \text{ and } Y \leq x)$

$$\overset{\text{ind.}}{=} P(X \leq x) \cdot P(Y \leq x)$$

$$= \left(1 - e^{-\lambda x}\right)^2 \qquad \text{for } x \geq 0$$

- The distribution function $F_{max(X,Y)}(x) = \left(1 - e^{-\lambda x}\right)^2$
- En $f_{max(X,Y)}(x) = \frac{d}{dx} F_{max(X,Y)}(x)$

$$= 2(1 - e^{-\lambda x}) \cdot \lambda e^{-\lambda x}, \text{ for x} \geq 0$$

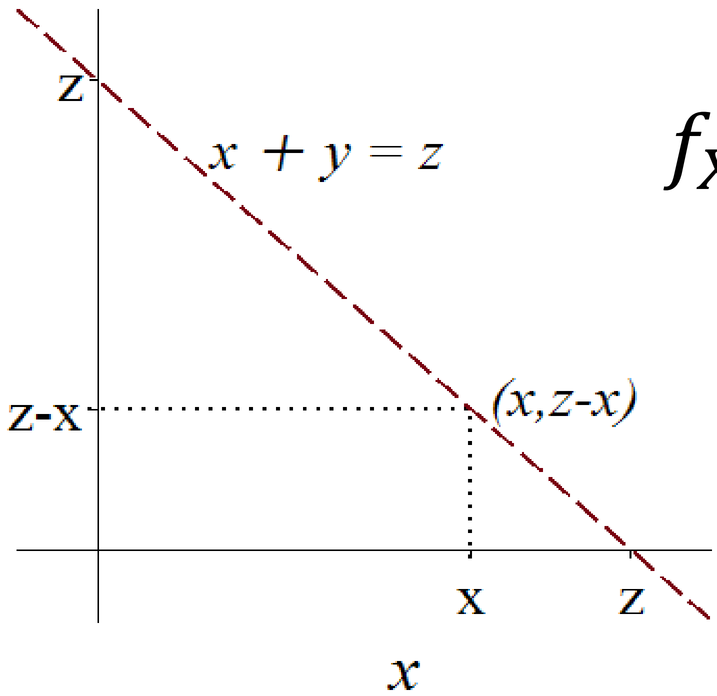
- **min**$(\boldsymbol{X},\boldsymbol{Y})$**:** $P(\min(X,Y) \geq m) = P(X \geq m \text{ and } Y \geq m)$

It follows: $f_{min(X,Y)}(m) = 2\lambda e^{-2\lambda m}$: $min(X,Y) \sim \boldsymbol{Exp}(\mathbf{2}\lambda)$

$X + Y$: de convolution integral

$$\mathrm{P}(X + Y = z) = \sum_x P(X = x)P(Y = z - x)$$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,dx$$



Example: X and Y are ind. and both $Exp(\lambda = 3)$, so:

$$f_{X+Y}(z) = \int_0^z 3e^{-3x} \cdot 3e^{-3(z-x)}dx$$

$$= \int_0^z 9e^{-3z}dx = 9e^{-3z} \cdot x\Big|_{x=0}^{x=z} = 9ze^{-3z} \;(z \geq 0)$$

The sum of normally distributed variables

Properties of a $N(\mu, \sigma^2)$-distributed X (chapter 6):

- Standardization: $\boldsymbol{Z = \frac{X - \mu}{\sigma} \sim N(0, 1)}$
- $\mathbf{Y} = \boldsymbol{aX + b \sim N\left(a\mu + b, a^2\sigma^2\right)}$

<u>Consequence **of the convolution integral**:</u>

If X and Y are independent and

$X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$,

then

$$\boldsymbol{X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)}$$

And

$$X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

Sample distributions

- A **random sample** from a normal distribution is a sequence of **independent** and $\boldsymbol{N(\mu, \sigma^2)}$**-distributed** $\boldsymbol{X_1, X_2, \ldots, X_n}$**.**
- Then for the **sum** we have:

$$\sum_{i=1}^{n} X_i = X_1 + \cdots + X_n \quad \sim \quad \boldsymbol{N(n\mu, n\sigma^2)}$$

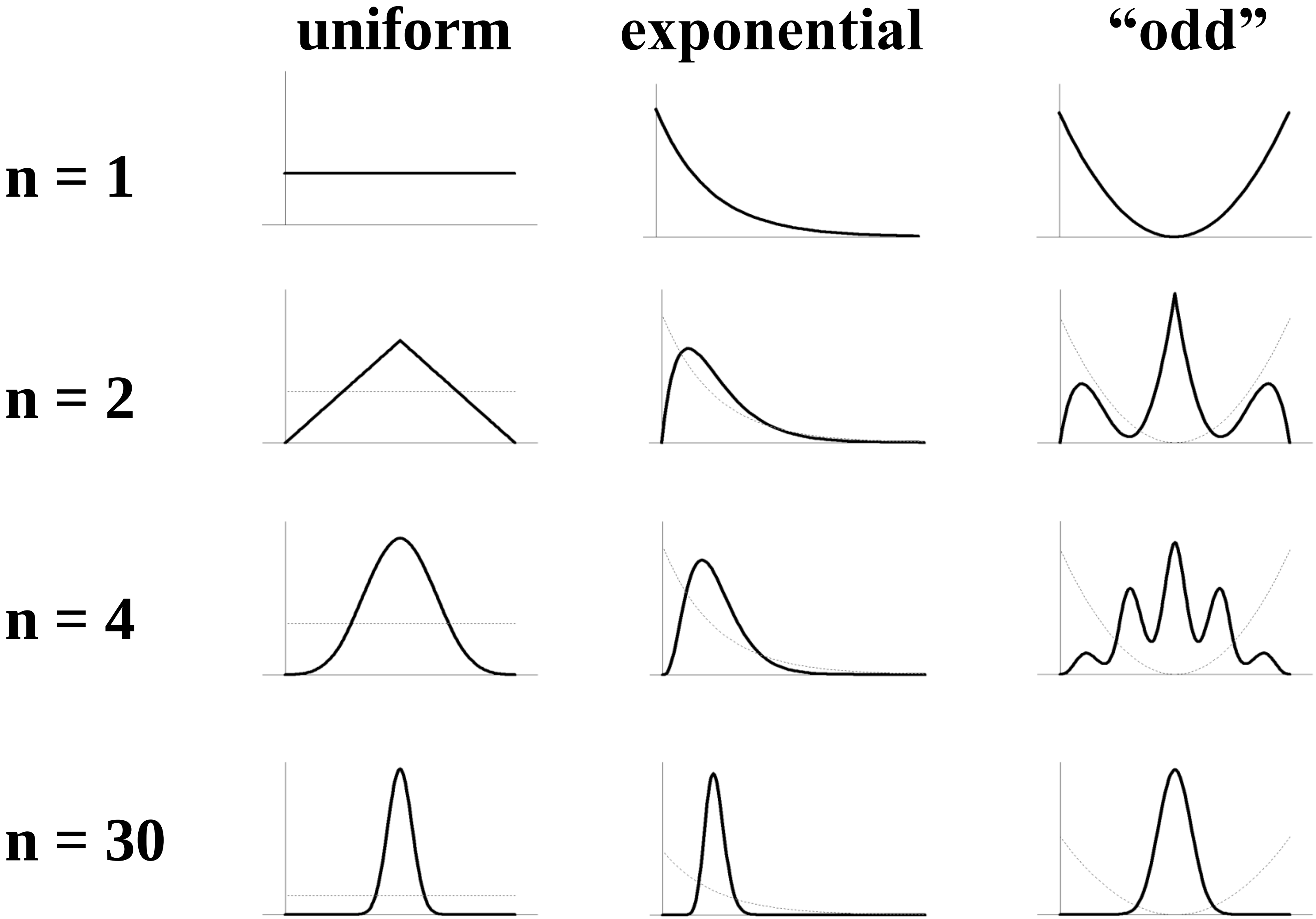
- And for **the sample mean:**

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \sim \quad \boldsymbol{N\left(\mu, \frac{\sigma^2}{n}\right)}$$

- So: $\sigma_{\overline{X}} = \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$

The distribution of the sample mean

if $X_1, \ldots, X_n$ are ind., all with the same distribution:



The Central Limit Theorem (CLT)

for independent and identically distributed $X_1, \ldots, Xn$ (**not normal**, but with expectation μ and variance σ^2):

$$\lim_{n\to\infty} P\left(\frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}} \le z\right) = \Phi(z)$$

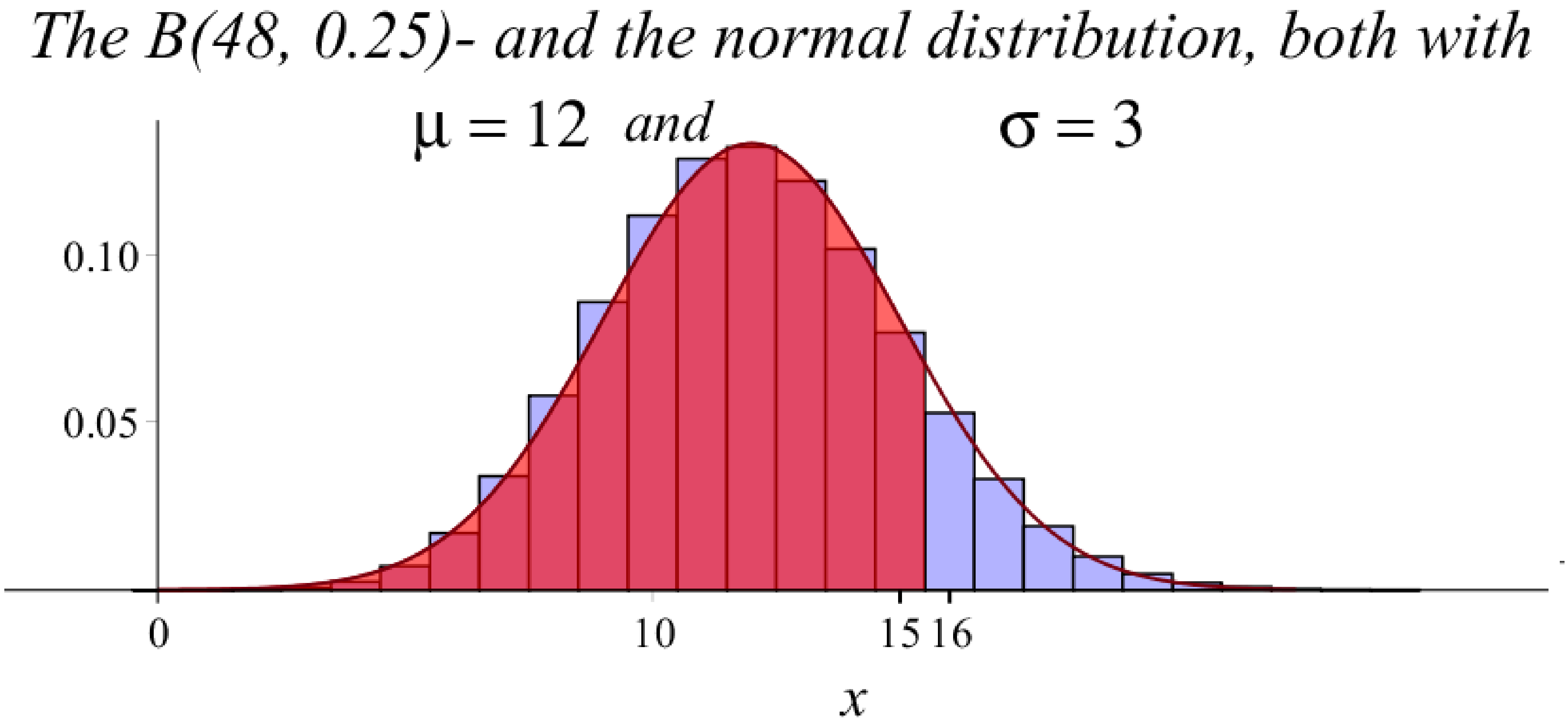
Note that $\frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}}$ is the result of standardization of $\sum X_i$

<u>Application</u>: for finite, **large *n* (rule of thumb: *n* > 25**), we have **approximately:**

$$\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2) \text{ and } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Approximation of $B(n, p)$ with $N(\mu, \sigma^2)$

$X \sim B(48, ¼)$ approximated by $Y \sim N(12, 9)$



The graphs illustrate: $\boldsymbol{P(X \leq 15) \approx P\left(Y \leq 15 + \frac{1}{2}\right)}$

Normal approximation of the binomial distribution

Approximate $\boldsymbol{X \sim B(n,p)}$ by $\boldsymbol{Y \sim N(np, np(1-p))}$

Consequence of the CLT: $\boldsymbol{X = \sum_{i=1}^{n} X_i}$**, where the** $\boldsymbol{X_i}$**`are independent alternatives**

Condition:

"n is sufficiently large and p not too close to 0 or 1"

rule of thumb: $\boldsymbol{n > 25, np > 5}$ **and** $\boldsymbol{n(1-p) > 5}$**.**

For this approximation **always apply continuity correction (c.c.):**

$$\boldsymbol{P(X \leq k) \overset{c.c.}{=} P(Y \leq k + 0.5)} \quad \text{and}$$

$$\boldsymbol{P(X < k) \overset{c.c.}{=} P(Y \leq k - 0.5)}$$