# VE401 Mid RC Pt.2

Wang Yangyang

UM-SJTU JI

2022 Spring

# Outline

1. **Continuous Random Variables**
   - Basics
   - Exponential, Gamma, Chi-Squared Distribution
   - Normal Distribution
   - Transformation of R.V.
   - Standardizing and Normal Approximation
   - Chebyshev's Inequality
2. **Multivariate Random Variables**
   - Discrete, Continuous Multivariate R.V.
   - Covariance and Correlation
   - Hypergeometric Distribution
   - Transformation of R.V. and Application
3. **Reliability**
   - Reliability and Hazard
   - Distribution and System

# Outline

# Continuous Random Variable: Def. and PDF

**Definition**

Let $S$ be a sample space. A *continuous random variable* is a map $X : S \to \mathbb{R}$ together with a function $f_X : \mathbb{R} \to \mathbb{R}$ with the properties that

1. $f_X(x) \geqslant 0$ for all $x \in \mathbb{R}$ and
2. $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$.

The integral of $f_X$ is interpreted as the probability that $X$ assumes values $X$ in a given range, i.e.,

$$P[a \leqslant X \leqslant b] = \int_a^b f_X(x)\mathrm{d}x$$

The function $f_X$ is called the *probability density function* of random variable $X$.

# Continuous Random Variable: CDF and Location

**Definition**

Let $(X, f_X)$ be a continuous random variable. The *cumulative distribution function* for $X$ is defined by $F_X : \mathbb{R} \to \mathbb{R}$,

$$F_X(x) := P[X \leqslant x] = \int_{-\infty}^{x} f_X(y)\mathrm{d}y$$

By the fundamental theorem of calculus, we can obtain the density function from $F_X$ by

$$f_X(x) = F_X'(x)$$

**Definition**

- The *median* $M_X$ is defined by $P[X \leqslant M_X] = 0.5$.
- The *mean* is given by $\mathrm{E}[X]$.
- The *mode* $x_0$, is the location of the maximum of $f_X$.

# Expectation, Variance and Moments

- Expectation.

$$\mathrm{E}[X] := \int_{\mathbb{R}} x \cdot f_X(x) \mathrm{d}x$$

- Variance.

$$\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$$

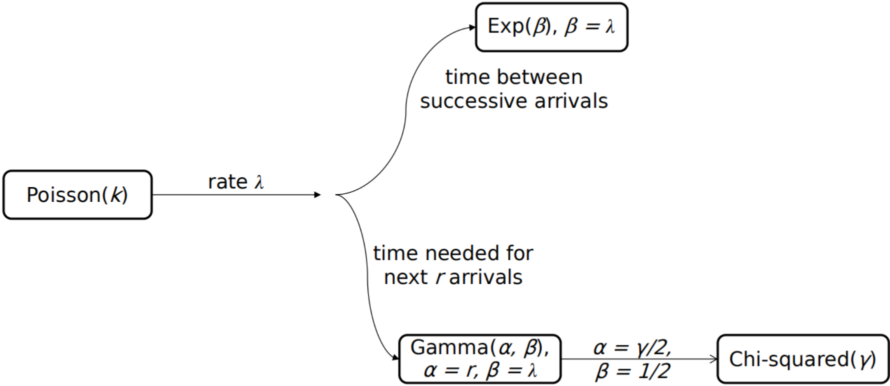
- M.G.F.

$$m_X(t) = \mathrm{E}\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} f_X(x) \mathrm{d}x$$

***Note:*** All previous properties about expectation, variance and M.G.F. hold for continuous random variables.

# Outline

# Connections



# Exponential Distribution

### Definition

A continuous random variable $(X, f_\beta)$ follows ***exponential distribution*** with parameter $\beta$ if the probability density function is defined by

$$f_\beta(X) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

### Interpretation

The time between successive arrivals of a Poisson process with rate $\lambda$ follows exponential distribution with parameter $\beta = \lambda$. (Recall $P[T > t] = e^{-\beta t}$.) ***Note.*** Memoryless property:

$$P[X > x + s \mid X > x] = P[X > s]$$

# Exponential Distribution

**Properties**

- Mean.

$$\mathrm{E}[X] = \frac{1}{\beta}$$

- Variance.

$$\mathrm{Var}[X] = \frac{1}{\beta^2}$$

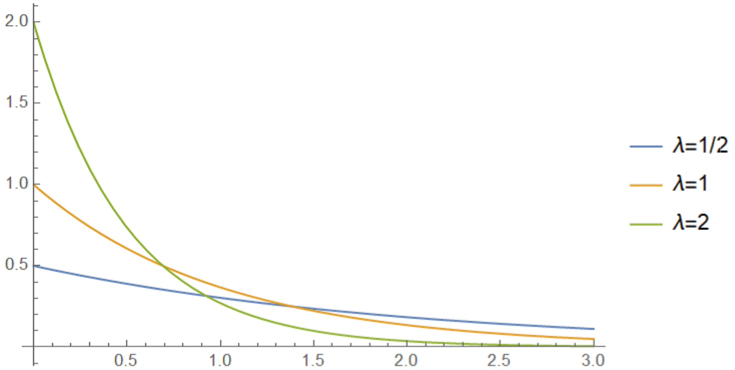
- M.G.F.

$$m_X : (-\infty, \beta) \to \mathbb{R}, \quad m_X(t) = \frac{1}{1 - t/\beta}$$

# Exponential Distribution

## Plot

```
Plot[Table[PDF[ExponentialDistribution[λ], x], {λ, {1/2, 1, 2}}]
// Evaluate, {x, 0, 3}, PlotRange -> All,
PlotLegends -> {"λ=1/2", "λ=1", "λ=2"}]
```



# Gamma Distribution

### Definition

Let $\alpha, \beta \in \mathbb{R}, \alpha, \beta > 0$. A continuous random variable $(X, f_{\alpha,\beta})$ follows a gamma distribution with parameters $\alpha$ and $\beta$ if the probability density function is given by

$$f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} \, \mathrm{d}z, \alpha > 0$ is the Euler gamma function.
Interpretation.

### Interpretation

The time needed for the next $r$ arrivals in a Poisson process with rate $\lambda$ follows a Gamma distribution with parameters $\alpha = r, \beta = \lambda$.

# Gamma Distribution

**Properties**

- Mean.

$$\mathrm{E}[X] = \frac{\alpha}{\beta}$$

- Variance.

$$\mathrm{Var}[X] = \frac{\alpha}{\beta^2}$$

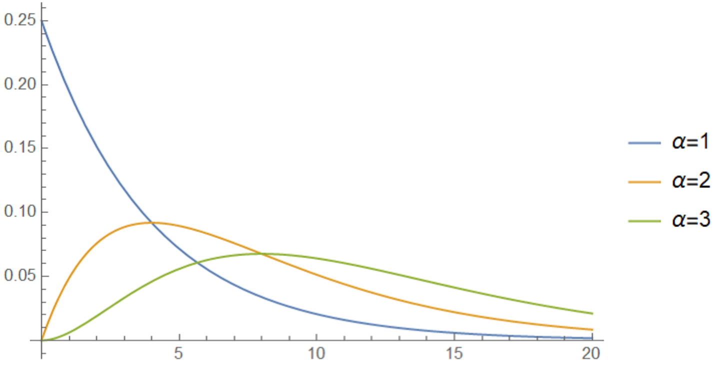
- M.G.F.

$$m_X : (-\infty, \beta) \to \mathbb{R}, \quad m_X(t) = \frac{1}{(1 - t/\beta)^{\alpha}}$$

# Gamma Distribution

## Plot

```
Plot[Table[PDF[GammaDistribution[α,4], x], {α, {1, 2, 3}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"α=1", "α=2", "α=3"}]
```

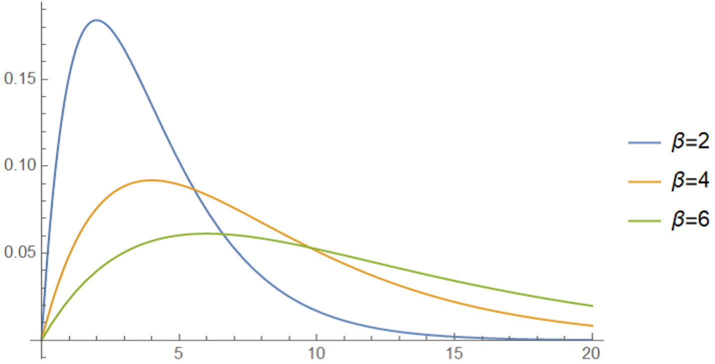


# Gamma Distribution

## Plot

What's Wrong? – FOR β, MMA IS DIFFERENT FROM THE LECTURE! It's $(1/2, 1/4, 1/6)$ for the lecture version.

```
Plot[Table[PDF[GammaDistribution[2,β], x], {β, {2, 4, 6}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"β=2", "β=4", "β=6"}]
```



# Chi-Squared Distribution

**Definition**

Definition. Let $\gamma \in \mathbb{N}$. A continuous random variable $\left(X_\gamma^2, f_X\right)$ follows a chi-squared distribution with $\gamma$ degrees of freedom if the probability density function is given by

$$f_\gamma(x) = \begin{cases} \frac{1}{2^{\gamma/2}\Gamma(\gamma/2)} x^{\gamma/2-1} e^{-x/2}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

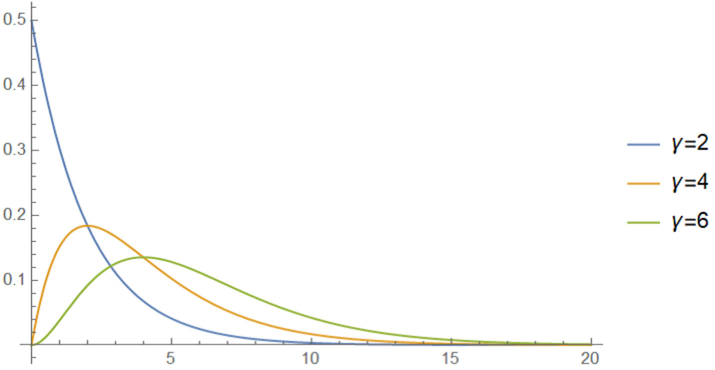
It is a gamma distribution with $\alpha = \gamma/2, \beta = 1/2$. Therefore,

$$\mathrm{E}\left[X_\gamma^2\right] = \gamma, \quad \mathrm{Var}\left[X_\gamma^2\right] = 2\gamma$$

# Chi-Squared Distribution

## Plot

```
Plot[Table[PDF[ChiSquareDistribution[γ], x], {γ, {2, 4, 6}}]
// Evaluate, {x, 0, 20}, PlotRange -> All,
PlotLegends -> {"γ=2", "γ=4", "γ=6"}]
```



# Outline

# Normal Distribution

## Definition

A continuous random variable $(X, f_{\mu,\sigma^2})$ has the *normal distribution* with mean $\mu \in \mathbb{R}$ and variance $\sigma^2$, $\sigma > 0$ if the probability density function is given by

$$f_{\mu,\sigma^2} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad x \in \mathbb{R}$$

## Useful Formula

Normal.

$$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$$

Gamma.

$$\int_0^{\infty} x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$$

# Normal Distribution

**Properties**

- Mean.

$$\mathrm{E}[X] = \mu$$

- Variance.

$$\mathrm{Var}[X] = \sigma^2$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \quad m_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

# Outline

# Transformation of Random Variables

**Theorem**

Let $X$ be a continuous random variable with density $f_X$. Let $Y = \varphi \circ X$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly *monotonic and differentiable*. The density for $Y$ is then given by

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|\frac{\mathrm{d}\varphi^{-1}(y)}{\mathrm{d}y}\right|,$$

for $y \in \operatorname{ran} \varphi$
and

$$f_Y(y) = 0,$$

for $y \notin \operatorname{ran} \varphi$.

What if not *monotonic and differentiable*? Consider *CDF*.

# Transformation of Random Variables

What if not ***monotonic and differentiable***? Consider ***CDF***.

**Example**

Consider the continuous random variable $X$ with density

$$f_X(x) = \frac{2/\pi}{e^{-x} + e^{x}}$$

for $x \in \mathbb{R}$
Find the density of the random variable $X^2$.

Things become easier using the observation that $f_X$ is an even function.

# Transformation of Random Variables

**Solution**

Note that the function $\varphi : \mathbb{R} \to \mathbb{R}_+ \cup \{0\}, \varphi(x) = x^2$ is not bijective, so we can't simply apply the theorem for transforming random variables from the lecture. Let $y > 0$. Then, using the fact that $f_X$ is even,

$$F_Y(y) = P[Y \leqslant y] = P\left[X^2 \leqslant y\right] = P[-\sqrt{y} \leqslant X \leqslant \sqrt{y}] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx$$

$$F_Y(y) = P[Y \leqslant y] = P\left[X^2 \leqslant y\right] = 2\int_0^{\sqrt{y}} f_X(x)dx$$

$$f_Y(y) = F_Y'(y) = 2f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \frac{2}{\pi\sqrt{y}} \frac{1}{e^{-\sqrt{y}} + e^{\sqrt{y}}} \quad y > 0$$

For $y \leqslant 0$ we have $F_Y(y) = P[Y \leqslant y] = P\left[X^2 \leqslant y\right] = 0$, so

$$f_Y(y) = 0 \quad y \leqslant 0.$$

# Outline

# Standardizing Normal Distribution

Suppose $X \sim \text{Normal}\left(\mu, \sigma^2\right)$. Then

$$Z = \frac{X - \mu}{\sigma} \sim \text{Normal}(0, 1)$$

where the normal distribution with mean $\mu$ and variance $\sigma^2$ is the *standard normal distribution*.

$\Phi$ is the cumulative distribution function for the standard normal distribution function.

$$\Phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

Calculate $P[X < a]$ by

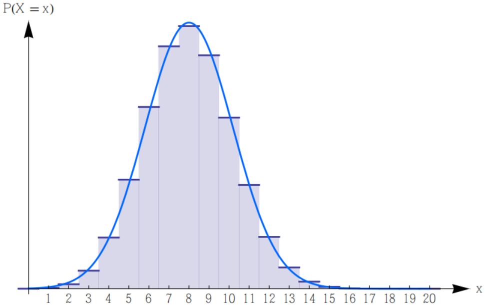
$$P[X < a] = P\left[\frac{X - \mu}{\sigma} < \frac{a - \mu}{\sigma}\right] = P\left[Z < \frac{a - \mu}{\sigma}\right] \quad = \Phi\left(\frac{a - \mu}{\sigma}\right)$$

# Theorem of De Moivre-Laplace

**Theorem**

Suppose $S_n$ is the number of successes in a sequence of $n$ i.i.d. Bernoulli trials with probability of success $0 < p < 1$. Then

$$\lim_{n\to\infty} P\left[a < \frac{X - np}{\sqrt{np(1-p)}} \leqslant b\right] = \frac{1}{2\pi}\int_a^b e^{-x^2/2}\,\mathrm{d}x$$



# Normal Approximation of Binomial Distribution

For $y = 0, \dots, n$

$$P[X \leqslant y] = \sum_{x=0}^{y} \binom{n}{x} p^x (1-p)^{n-x} \approx \Phi\left(\frac{y + 1/2 - np}{\sqrt{np(1-p)}}\right)$$

where we require that

$$np > 5 \quad \text{if } p \leqslant \frac{1}{2} \quad \text{or} \quad n(1-p) > 5 \quad \text{if } p > \frac{1}{2}$$

This additional term $1/2$ is known as the ***half-unit correction*** for the normal approximation to the ***cumulative*** binomial distribution function.

# Outline

# Chebyshev's Inequality and Law of Large Number

**Theorem**

Let $X$ be a random variable, then for $k \in \mathbb{N} \backslash \{0\}$ and $c > 0$,

$$P[|X| \geqslant c] \leqslant \frac{\mathrm{E}\left[|X|^k\right]}{c^k}$$

As another version of this inequality, suppose $X$ has mean $\mu$ and standard deviation $\sigma$, and let $m > 0$,

$$P[|X - \mu| \geqslant m\sigma] \leqslant \frac{1}{m^2}$$

**Weak Law of Large Number**

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean $\mu$ and variance $\sigma^2$. Then for any $\varepsilon > 0$,

$$P\left[\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| \geqslant \varepsilon\right] \xrightarrow{n \to \infty} 0$$

# Central Limit Theorem

### Theorem

Central Limit Theorem. Let $(X_i)$ be a sequence of independent, but not necessarily identical random variables whose third moments exist and satisfy a certain technical condition. Let

$$Y_n = X_1 + \cdots + X_n$$

Then for any $z \in \mathbb{R}$

$$P\left[\frac{Y_n - \mathrm{E}[Y_n]}{\sqrt{\mathrm{Var}[Y_n]}} \leqslant z\right] \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

### Interpretation

Lyapunov's Central Limit theorem is at the core of the belief by experimentalists that "random error" may be described by the normal distribution.

# Outline

# Discrete Multivariate Random Variable

**Definition**

Let $S$ be a sample space and $\Omega$ a countable subset of $\mathbb{R}^n$. A *discrete multivariate random variable* is a map

$$\mathrm{X} : S \to \Omega$$

together with a function $f_{\mathrm{X}} : \Omega \to \mathbb{R}$ with the properties that

1. $f_{\mathrm{X}}(x) \geqslant 0$ for all $x = (x_1, \ldots, x_n) \in \Omega$ and
2. $\sum_{x \in \Omega} f_{\mathrm{X}}(x) = 1$,

where $f_{\mathrm{X}}$ is the *joint density function* of the random variable X.

# Continuous Multivariate Random Variable

**Definition**

Let $S$ be a sample space. A *continuous multivariate random variable* is a map

$$\mathrm{X} : S \to \mathbb{R}^n$$

together with a function $f_{\mathrm{X}} : \mathbb{R}^n \to \mathbb{R}$ with the properties that

1. $f_{\mathrm{X}}(x) \geqslant 0$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and
2. $\int_{\mathbb{R}^n} f_{\mathrm{X}}(x) = 1$,

where $f_{\mathrm{X}}$ is the *joint density function* of the random variable X.

# Multivariate Random Variable

- *Marginal density* $f_{X_k}$ for $X_k, k = 1, \ldots, n$ :

$$f_{X_k}(x_k) = \int_{\mathbb{R}^{n-1}} f_X(x_1, \ldots, x_n) \, \mathrm{d}x_1 \ldots \mathrm{d}x_{k-1} \, \mathrm{d}x_{k+1} \ldots \mathrm{d}x_n.$$

- *Independent* multivariate random variables:

$$f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

- *Conditional density* of $X_1$ conditioned on $X_2$ :

$$f_{X_1 | X_2}(x_1) := \frac{f_{X_1 X_2}(x_1, x_2)}{f_{X_2}(x_2)} \quad \text{whenever } f_{X_2}(x_2) > 0.$$

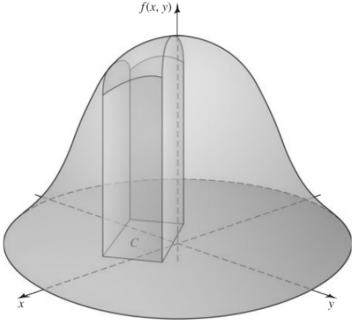
- *Joint cumulative distribution function* (CDF) is then given by

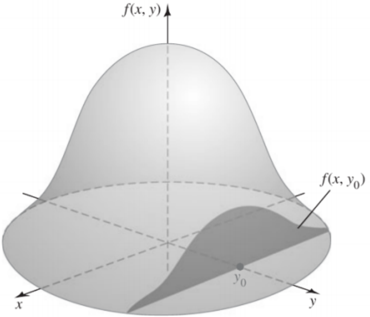
$$P[X_1 \leqslant a_1, \ldots, X_n \leqslant a_n] = \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} f_X(x) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

# Continuous Multivariate Random Variable

**Visualization**

Joint probability density function $f_{XY}(x, y)$ (left) conditional density function $f_{X|Y}(x \mid y_0)$ (right).



# Outline

# Expectation

- Discrete.

$$\mathrm{E}\left[X_k\right] = \sum_{x_k} x_k f_{X_k}\left(x_k\right) = \sum_{x \in \Omega} x_k f_X(x)$$

and for continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{E}[\varphi \circ X] = \sum_{x \in \Omega} \varphi(x) f_X(x)$$

- Continuous.

$$\mathrm{E}\left[X_k\right] = \int_{\mathbb{R}} x_k f_{X_k}\left(x_k\right) \mathrm{d}x_k = \int_{\mathbb{R}^n} x_k f_X(x) \mathrm{d}x$$

and for continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{E}[\varphi \circ X] = \int_{\mathbb{R}^n} \varphi(x) f_X(x) \mathrm{d}x.$$

# Covariance

### Definition

Definition. For a multivariate random variable X, the ***covariance matrix*** Var[X] is given by

$$\begin{pmatrix} \operatorname{Var}[X_1] & \operatorname{Cov}[X_1, X_2] & \cdots & \operatorname{Cov}[X_1, X_n] \\ \operatorname{Cov}[X_1, X_2] & \operatorname{Var}[X_2] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \operatorname{Cov}[X_{n-1}, X_n] \\ \operatorname{Cov}[X_1, X_n] & \cdots & \operatorname{Cov}[X_{n-1}, X_n] & \operatorname{Var}[X_n] \end{pmatrix}$$

where the ***covariance*** of $(X_i, X_j)$ is given by

$$\operatorname{Cov}[X_i, X_j] = \mathrm{E}\left[(X_i - \mu_{X_i})\left(X_j - \mu_{X_j}\right)\right] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\,\mathrm{E}[X_j]$$

and

$$\operatorname{Var}[\mathrm{C}\mathrm{X}] = \mathrm{C}\operatorname{Var}[\mathrm{X}]\mathrm{C}^T, \quad \mathrm{C} \in \operatorname{Mat}(n \times n; \mathbb{R})$$

# Covariance

## Properties

Let $X, X_1, \ldots, X_n$, $Y$ and $Z$ be random variables.

- $X$ and $Y$ are independent $\Rightarrow \text{Cov}[X, Y] = 0$.
  The Converse is not True!
- $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\,\text{Cov}[X, Y]$, and more generally,

  $$\text{Var}\left[X_1 + \cdots + X_n\right] = \text{Var}\left[X_1\right] + \cdots + \text{Var}\left[X_n\right] + 2\sum_{i<j} \text{Cov}\left[X_i, X_j\right]$$

  if $\text{Var}\left[X_i\right] < \infty$ for $i = 1, \ldots, n$
- $\text{Cov}[X, Y + Z] = \text{Cov}[X, Y] + \text{Cov}[X, Z]$
  $\text{Cov}[X, Y - Z] = \text{Cov}[X, Y] - \text{Cov}[X, Z]$
- $\text{Cov}[X, X] = \text{Var}[X]$

# Correlation

**Definition**

The *Pearson coefficient of correlation* of random variables $X$ and $Y$ is given by

$$\rho_{XY} := \frac{\operatorname{Cov}[X, Y]}{\sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]}}$$

Instead of independence, the correlation coefficient actually measures the the extent to which $X$ and $Y$ are *linearly* dependent, which is not the only way of being dependent.

**Properties**

- $-1 \leqslant \rho_{XY} \leqslant 1$,
- $|\rho_{XY}| = 1$ iff there exist $\beta_0, \beta_1 \in \mathbb{R}$ such that

$$Y = \beta_0 + \beta_1 X$$

# Correlation

### Note

- Uncorrelated does not mean independent,
- Correlation coefficient only measures linear relationships.

### Example

Two random variables $X$ and $Y$.$X$ follows a uniform distribution $U(-1,1)$ and $Y = X^2$. Find $\mathrm{Cov}(X, Y)$.

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{Cov}\left(X, X^2\right) \\
&= \mathrm{E}\left((X - \mathrm{E}(X))\left(X^2 - \mathrm{E}\left(X^2\right)\right)\right) \\
&= \mathrm{E}\left(X^3 - X^2\mathrm{E}(X) - X\mathrm{E}\left(X^2\right) + \mathrm{E}(X)\mathrm{E}\left(X^2\right)\right) \\
&= \mathrm{E}\left(X^3\right) - \mathrm{E}\left(X^2\right)\mathrm{E}(X) - \mathrm{E}(X)\mathrm{E}\left(X^2\right) + \mathrm{E}(X)\mathrm{E}\left(X^2\right) \\
&= \int_{-1}^{1} \frac{1}{2}x^3 \, \mathrm{d}x - \int_{-1}^{1} \frac{1}{2}x^2 \, \mathrm{d}x \cdot \int_{-1}^{1} \frac{1}{2}x \, \mathrm{d}x = 0
\end{aligned}$$

# The Fisher Transformation

**Definition**

Let $\widetilde{X}$ and $\widetilde{Y}$ be standardized random variables of $X$ and $Y$, then the *Fisher transformation* of $\rho_{XY}$ is given by

$$\ln\left(\sqrt{\frac{\mathrm{Var}[\widetilde{X}+\widetilde{Y}]}{\mathrm{Var}[\widetilde{X}-\widetilde{Y}]}}\right) = \frac{1}{2}\ln\left(\frac{1+\rho_{XY}}{1-\rho_{XY}}\right) = \mathrm{Arctanh}\,(\rho_{XY}) \in \mathbb{R}$$

We say that $X$ and $Y$ are

- *positively correlated* if $\rho_{XY} > 0$, and
- *negatively correlated* if $\rho_{XY} < 0$.

# The Bivariate Normal Distribution

The density function of ***Bivariate Normal Distribution***:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$

- $-1 < \rho < 1$
- $\mu_X = \mathrm{E}[X]$, $\sigma_X^2 = \operatorname{Var} X$ (and similarly for $Y$ ).
- $\rho = \rho_{XY}$ is indeed the correlation coefficient of $X$ and $Y$.
- $X$ and $Y$ are independent $\iff \rho = 0$

# Outline

# Hypergeometric Distribution

## Definition

A random variable $(X, f_X)$ with parameters $N, n, r \in \mathbb{N}\backslash\{0\}$ where $r, n \leqslant N$ and $n < \min\{r, N-r\}$ has a ***hypergeometric distribution*** if the density function is given by

$$f_X(x) = \frac{\begin{pmatrix} r \\ x \end{pmatrix} \begin{pmatrix} N-r \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

## Interpretation

- $f_X(x)$ is the probability of getting $x$ red balls in drawing $n$ balls from a box containing $N$ balls, where $r$ of them are red.
- This can be formulated as obtaining $x$ successes in $n$ identical but not independent Bernoulli trials, each with probability of success $\frac{r}{N}$.

# Hypergeometric Distribution

**Property**

- Expectation

$$\mathrm{E}[X] = \mathrm{E}\left[X_1 + \cdots + X_n\right] = n\frac{r}{N}$$

- Variance

$$\begin{aligned}
\mathrm{Var}[X] &= \mathrm{Var}\left[X_1 + \cdots + X_n\right] \\
&= \mathrm{Var}\left[X_1\right] + \cdots + \mathrm{Var}\left[X_n\right] + 2\sum_{i<j} \mathrm{Cov}\left[X_i, X_j\right] \\
&= n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}
\end{aligned}$$

The *binomial distribution* may be used to approximate the hypergeometric distribution if $n/N$ is small (less than 0.05).

# Outline

# Transformation of Random Variables

**Theorem**

Let $(\boldsymbol{X}, f_{\boldsymbol{X}})$ be a continuous multivariate random variable and let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable, bijective map with inverse $\varphi^{-1}$. Then $\boldsymbol{Y} = \varphi \circ \boldsymbol{X}$ is a continuous multivariate random variable with density

$$f_Y(y) = f_X \circ \varphi^{-1}(y) \cdot \left|\det D\varphi^{-1}(y)\right|,$$

where $D\varphi^{-1}$ is the Jacobian of $\varphi^{-1}$.

- $f_Y(y) = 0$ for $y \notin \operatorname{ran} \varphi$.

# Quotient of Normal: Cauchy

**Lemma**

Let $((X, Y), f_{XY})$ be a continuous bivariate random variable. Let $U = X/Y$. Then the density $f_U$ of $U$ is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_{XY}(uv, v) \cdot |v| dv.$$

**Theorem**

Suppose that random variables $X$ and $Y$ are independent and that each follows the ***standard normal distribution***. Then $U = X/Y$ has the ***Cauchy distribution*** with probability density function given by

$$f_U(u) = \frac{1}{\pi(1+u^2)}, \quad u \in \mathbb{R}.$$

# Quotient of Normal: Cauchy

**Proof**

Let $V = Y$, excluding $Y = 0$, the transformation from $(X, Y)$ to $(U, V)$ is one-to-one. Then $X = UV$, $Y = V$ and

$$J = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = v$$

Then the joint density function is given by

$$f_{UV}(u, v) = f_{XY}(uv, v)|v| = \frac{|v|}{2\pi} \exp\left(-\frac{1}{2}\left(u^2+1\right)v^2\right)$$

Then the marginal of $U$ is calculated as

$$f_U(u) = \int_{-\infty}^{\infty} f_{Uv}(u, v)\mathrm{d}v = \frac{1}{\pi\left(u^2+1\right)}, \quad u \in \mathbb{R}.$$

# Root Sum of Normal Square: Chi

**Definition**

$\chi_n$ is a *chi random variable* with $n$ *degrees of freedom*,

$$\chi_n = \sqrt{\sum_{i=1}^{n} Z_i^2}$$

where $Z_1, \ldots, Z_n$ are *independent standard normal* random variables.

$$f_{\chi_n}(y) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} y^{n-1} e^{-y^2/2} \quad (y > 0)$$

**Interpretation**

A chi random variable represents the sum of the root squares (distance) of independent standard normal variables.

# Sum of Normal Square: Chi-Squared

**Definition**

$\chi_n^2$ is a *chi-squared random variable* with $n$ *degrees of freedom*,

$$\chi_n^2 = \sum_{i=1}^{n} Z_i^2$$

where $Z_1, \ldots, Z_n$ are *independent standard normal* random variables.

$$f_{\chi_n^2}(y) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} y^{n/2-1} e^{-y/2} \quad (y > 0)$$

**Interpretation**

A chi-squared random variable represents the sum of the squares of independent standard normal variables.

# Sum of Normal: Normal

**Theorem**

If the random variables $X_1, \ldots, X_k$ are independent and if $X_i$ follows *normal distribution* with mean $\mu_i$ and variance $\sigma_i^2$, where $i = 1, \ldots, k$, then $X = X_1 + \cdots + X_k$ follows normal distribution with

$$\mu = \mu_1 + \cdots + \mu_k, \quad \sigma^2 = \sigma_1^2 + \cdots + \sigma_k^2.$$

**Theorem**

Let $\chi^2_{\gamma_1}, \ldots, \chi^2_{\gamma_n}$ be $n$ independent random variables following *chi-squared distributions* with $\gamma_1, \ldots, \gamma_n$ degrees of freedom, respectively. Then

$$\chi^2_\alpha := \sum_{k=1}^{n} \chi^2_{\gamma_k}$$

is a *chi-squared random variable* with $\alpha = \sum_{k=1}^{n} \gamma_k$ degrees of freedom.

# Outline

# Reliability

## Definition

Suppose a unit $A$ fails ***randomly***, and we describe the time it fails by the continuous random variable $T_A$.
The density of $T_A$ is called the ***failure density*** $f_A$. The cumulative distribution function of $T_A$ is denoted by $F_A$. We note that

$$\begin{aligned} f_A(t) &= \lim_{\Delta t \to 0} \frac{P[t \leqslant T \leqslant t + \Delta t]}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{F_A(t + \Delta t) - F_A(t)}{\Delta t} \end{aligned}$$

The ***reliability function*** $R_A$ gives the probability that $A$ is working at time $t \geqslant 0$

$$\begin{aligned} R_A(0) &= 1 \\ R_A(t) &= 1 - P[\text{ component A fails before time } t] \\ &= 1 - F_A(t) \end{aligned}$$

# Hazard

**Definition**

*Hazard rate* $\rho_A$ defined by

$$\rho_A(t) = \lim_{\Delta t \to 0} \frac{P[t \leqslant T \leqslant t + \Delta t \mid t \leqslant T]}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{P[t \leqslant T \leqslant t + \Delta t]}{P[T \geqslant t] \cdot \Delta t} = \frac{f_A(t)}{R_A(t)}$$

**Interpretation**

- If $\rho$ is decreasing, then as time goes by a failure is more likely to occur earlier in the time interval.
- If $\rho$ is steady, a failure tends to occur during this period due mainly to random factors.
- If $\rho$ is increasing, then as time goes by a failure is more likely to occur.

# Outline

# Exponential Distribution

- Density function. $\beta > 0$ is a parameter,

$$f(x) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

- Mean.

$$\mu = \frac{1}{\beta}$$

- Variance.

$$\sigma^2 = \frac{1}{\beta^2}$$

- Reliability features.

$$\rho(t) = \beta$$

$$R(t) = e^{-\beta t}, f(t) = \rho(t)R(t) = \beta e^{-\beta t}.$$

# Weibull Distribution

- Density function. $\alpha, \beta > 0$ are parameters,

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

- Mean.

$$\mu = \alpha^{-1/\beta}\Gamma(1 + 1/\beta)$$

- Variance.

$$\sigma^2 = \alpha^{-2/\beta}\Gamma(1 + 2/\beta) - \mu^2$$

- Reliability features.

$$\rho(t) = \alpha\beta t^{\beta-1}$$

$$R(t) = e^{-\alpha t^{\beta}}, f(t) = \rho(t)R(t) = \alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}$$

# System

$R_i$ is the reliability of the ith component, then

1. reliability of a series system with $k$ components

$$R_s(t) = \prod_{i=1}^{k} R_i(t)$$

2. reliability of a parallel system with $k$ components

$$R_p(t) = 1 - P[\text{all components fail before t}] = 1 - \prod_{i=1}^{k} (1 - R_i(t))$$

# End

Credit to Zhanpeng Zhou (TA of SP21)
Credit to Fan Zhang (TA of SU21)
Credit to Jiawen Fan (TA of SP21)
Credit to Zhenghao Gu (TA of SP20)
`https://www.johndcook.com/blog/distribution_chart/`