# Chapter 1: Elements of Probability

## 1. Notation

Intersection: '$\cap$', '$\cdot$', "and"

Union: '$\cup$', '$+$', "or"

Complement of $A$: $\overline{A}$ or $A^c$ or $A^*$

## 2. Probability

Probability of an event $A$

$$\mathrm{P}(A) = \frac{\text{number of elements in } A}{\text{number of elements in } \Omega} = \frac{|A|}{|\Omega|}$$

Probability of a complement event $A^c$

$$\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$$

## 3. Properties

**Table 1:** *Properties of probabilities.*

| **Name** | **Property** |
|---|---|
| $A$ or $B$ | $\mathrm{P}(A \cup B) = \mathrm{P}(B \cup A) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ |
| $A$ and $B$ | $\mathrm{P}(A \cap B) = \mathrm{P}(B \cap A) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B)$ |
| $A$ exclude $B$ | $\mathrm{P}(A \setminus B) = \mathrm{P}(A) - \mathrm{P}(A \cap B)$ |
| Independence | $\mathrm{P}(AB) = \mathrm{P}(A)\mathrm{P}(B)$ |
| De Morgan's laws | $\mathrm{P}(\overline{A \cdot B}) = \mathrm{P}(\overline{A} + \overline{B})$ <br> $\mathrm{P}(\overline{A + B}) = \mathrm{P}(\overline{A} \cdot \overline{B})$ |
| $A$ given $B$ | $\mathrm{P}(A\|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ |
| Conditional complement | $\mathrm{P}(\overline{A}\|B) = 1 - \mathrm{P}(A\|B)$ |
| Changing order | $\mathrm{P}(A\|B) = \dfrac{\mathrm{P}(B\|A)P(A)}{\mathrm{P}(B)}$ |
| Total $B$ | $\mathrm{P}(B) = \mathrm{P}(B\|A)P(A) + \mathrm{P}(B\|A^c)\mathrm{P}(A^c)$ |

*(Nguyên tắc đếm: Trường hợp là cộng, đi tiếp là nhân)*

## Chapter 2: Random Variables (RV)

**Note:**

- cdf: cumulative distribution function.
- pmf: probability mass function.
- pdf: probability density function.
- Expectation (Expected value or Mean): $\mu = \mathrm{E}[X]$.
- Variance: $\sigma^2 = \mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}^2[X]$.
- Standard deviation: $\sigma$.

### 1. Cumulative Distribution Function

A cumulative distribution function (c.d.f or cdf) of a random variable $X$, evaluated at $x$, is the probability that $X$ will take a value less than or equal to $x$, or

$$F(x) = \mathrm{P}(X \leq x)$$

If a random variable $X$ is absolutely continuous univariate distributions then the pdf of $X$ is

$$f(x) = \frac{d}{dx}F(x)$$

### 2. Probability Mass Function

A probability mass function (p.m.f or pmf) is a function which **gives** us **directly probability** of a **Discrete RV** at a certain value. Denote that

$$\mathrm{P}(X = x_i) = p(x_i)$$

Normalized condition

$$\sum_{X} \mathrm{P}(X = x_i) = p(x_1) + p(x_2) + \cdots = 1$$

Pmf to cdf

$$F(x) = \mathrm{P}(X \leq x) = \sum_{x_i \leq x} p(x_i)$$

### 3. Probability Density Function

A probability density function (p.d.f or pdf) is a function that **gives** us **density** of a **Continuous RV**. Denote that $f(x)$ or $p(x)$.

Normalized condition

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Pdf to cdf

$$F(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f(\tau)d\tau$$

## 4. Expectation and Variance of RV

### 4. 1. Definition

**Table 2:** *Definition of Expectation and Variance.*

| | **Discrete RV** | **Continuous RV** |
|---|---|---|
| **Expectation** | $\mathrm{E}[X]=\sum_{X} X_i p(x_i)$ | $\mathrm{E}[X]=\int_X xp(x)dx$ |
| **Variance** | $\mathrm{Var}[X]=\mathrm{E}[X^2]-\mathrm{E}^2[X]$ | $\mathrm{Var}[X]=\mathrm{E}[X^2]-\mathrm{E}^2[X]$ |

### 4. 2. Properties

**Table 3:** *Properties of expectation.*

| **Name** | **Property** |
|---|---|
| Constant | $\mathrm{E}[c]=c$ |
| Linearity | $\mathrm{E}[aX+bY]=a\mathrm{E}[X]+b\mathrm{E}[Y]$ |
| Independent | $\mathrm{E}[XY]=\mathrm{E}[X]\cdot\mathrm{E}[Y]$ |

**Table 4:** *Properties of variance.*

| **Name** | **Property** |
|---|---|
| Constant | $\mathrm{Var}[c]=0$ |
| Invariant | $\mathrm{Var}[X+c]=\mathrm{Var}[X]$ |
| Non-Linearity | $\mathrm{Var}[aX+c]=a^2\mathrm{Var}[X]$ |
| Independent | $\mathrm{Var}[X+Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]$ |

## 5. Special Discrete RV

### 5. 1. Bernoulli RV

Discrete RV $X$ is called Bernoulli RV with parameter $p$ if its pmf is

$$X\colon \begin{cases}\mathrm{P}(X=0)=1-p\\ \mathrm{P}(X=1)=p\end{cases}$$

Denote: $X\sim\mathrm{Ber}(p)$.

Property: $\mathrm{E}[X]=p;\ \mathrm{Var}[X]=p(1-p)$.

### 5. 2. Geometric RV

Discrete RV with $X$ is number of independent trials until the first success, each is $\mathrm{Ber}(p)$ is called Geometric RV.

Denote: $X\sim\mathrm{Geo}(p)$.

Property: $\mathrm{E}[X]=1/p\,;\ \mathrm{Var}[X]=(1-p)/p^2$.

The pmf of Geometric RV

$$\mathrm{P}(X=i)=(1-p)^{i-1}p,\qquad i\geq 1$$

### 5. 3. Binomial RV

Discrete RV with $X$ is number of successes in $n$ independent trials, each is $\mathrm{Ber}(p)$ is called Binomial RV.

Denote: $X \sim \mathrm{Bino}(n, p)$.

Property: $\mathrm{E}[X] = np$; $\mathrm{Var}[X] = np(1-p)$.

Probability of event that $X = i$

$$\mathrm{P}(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

Probability of event that $a \leq X \leq b$

$$\mathrm{P}(a \leq X \leq b) = \sum_{i=a}^{b} \binom{n}{i} p^i (1-p)^{n-i}$$

### 5. 4. Poisson RV

Discrete RV $X$ is called Poisson RV with parameter $\lambda$ if the pmf is (Poisson RV can be considered as Binomial RV with large $n$ and small $p$ and $\lambda = np$)

$$\mathrm{P}(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

Denote: $X \sim \mathrm{Poisson}(\lambda)$.

Property: $\mathrm{E}[X] = \lambda$; $\mathrm{Var}[X] = \lambda$.

## 6. Special Continuous RV

### 6. 1. Uniform RV

Continuous RV $X$ is uniform RV on $[a, b]$ if its pdf is

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

Denote: $X \sim \mathrm{Uni}[a, b]$.

Property: $\mathrm{E}[X] = (a+b)/2$ ; $\mathrm{Var}[X] = (b-a)^2/12$.

### 6. 2. Exponential RV

Continuous RV $X$ is called exponential RV with parameter $\lambda$ if its pdf is

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

Denote: $X \sim \varepsilon(\lambda)$.

Property: $\mathrm{E}[X] = 1/\lambda$ ; $\mathrm{Var}[X] = 1/\lambda^2$.

This random variable is used for modeling arrival time (or date) of some event.

### 6. 3. Normal RV

Continuous RV $X$ is normally distributed with parameters $\mu$ and $\sigma^2$ if its pdf is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Denote: $X \sim \mathcal{N}(\mu, \sigma^2)$.

Property: $\mathrm{E}[X] = \mu$; $\mathrm{Var}[X] = \sigma^2$.

Probabilities

1. $\mathrm{P}(X < a) = \mathrm{P}\left(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right) = \mathrm{P}(Z < z_0)$

2. $\mathrm{P}(a \le X < b) = \mathrm{P}(z_1 \le Z < z_2) = \mathrm{P}(Z < z_2) - \mathrm{P}(Z < z_1)$

### 6. 3. 1. Calculating $z$-value

1. For FX570, VinaCal II:

- Go to statistics mode: MODE → 3 → AC.
- $\mathrm{P}(Z < z_0)$: $\mathrm{P}(z_0)$, Shift → 1 → 5 → 1.
- $\mathrm{P}(Z > z_0)$: $\mathrm{R}(z_0)$, Shift → 1 → 5 → 3.

2. For FX580:

- Go to statistics mode: MODE → 6 → AC.
- OPTN → ⇩ → “4: Norm Dist”.
- $\mathrm{P}(Z < z_0)$: $\mathrm{P}(z_0)$, Press 1.
- $\mathrm{P}(Z > z_0)$: $\mathrm{R}(z_0)$, Press 3.

### 6. 3. 2. Sum of Normal RV $X_i \sim \mathcal{N}(\mu, \sigma^2)$:

$$\mathrm{E}[X] = \sum_n \mathrm{E}[X_i] = n\mu$$

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

### 6. 3. 3. Average of Normal RV $X_i \sim \mathcal{N}(\mu, \sigma^2)$:

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

$$\mathrm{Var}[\bar{X}] = \frac{1}{n^2}\sum_n \mathrm{Var}[X_i] \ = \frac{\sigma^2}{n}$$

The above properties for sum and average is only applied for identical independent normal random variables $X_i, i = 0,1, \dots, n$. If $X_i$ are difference only the first equal sign are valid.

### 6. 3. 4. Chebyshev's Inequality

Let $X$ be a random variable with finite expected value $\mu$ and finite non-zero variance $\sigma^2$. Then for any real number $k > 0$,

$$\mathrm{P}(|X-\mu| \ge k\sigma) \ge \frac{1}{k^2}$$

# Chapter 3: Statistic

## 1. Introduction

**Table 5:** *Statistic parameters.*

| | Sample | Population |
|---|---|---|
| Mean | $\overline{X}$ | $\mu$ |
| Variance | $s^2$ | $\sigma^2$ |
| Standard deviation | $s$ | $\sigma$ |

Data Processing for Single Row/Column Data:

1. For FX570, VinaCal II:
   - Insert data: MODE → 3 → 1.
   - $\bar{X}$: Shift → 1 → 4 → 2.
   - $s$: Shift → 1 → 4 → 4.
2. For FX580:
   - Insert data: MODE → 6 → 1.
   - OPTN → 3.

## 2. Linear Regression

### 2. 1. Cartesian Data ($\boldsymbol{Y_i = \beta_0 + \beta_1 X_i}$)

$$\bar{X}, \bar{Y} \to S_{XY}, S_{XX}, S_{YY}, SS \to \beta_0, \beta_1$$

$$S_{XY} = \sum X_i Y_i - n\bar{X}.\bar{Y};\;\; S_{XX} = \sum X_i^2 - n\bar{X}^2;\;\; S_{YY} = \sum Y_i^2 - n\bar{Y}^2$$

$$SS = \frac{S_{XX}.S_{YY} - S_{XY}^2}{S_{XX}}$$

$$\beta_1 = \frac{S_{XY}}{S_{XX}};\;\; \beta_0 = \bar{Y} - \beta_1 \bar{X}$$

### 2. 2. Modeling Linear Regression

**Table 6:** *Instructions for modeling non-linear model.*

| Function Form | Transformation | Linear Regression |
|---|---|---|
| Exponential: $y = \beta_0 e^{\beta_1 x}$ | $\hat{y} = \ln y$ | $\hat{y} = \ln \beta_0 + \beta_1 x$ |
| Power: $y = \beta_0 x^{\beta_1}$ | $\hat{y} = \log y\,;\;\; \hat{x} = \log x$ | $\hat{y} = \log \beta_0 + \beta_1 \hat{x}$ |
| Reciprocal: $y = \beta_0 + \beta_1 \dfrac{1}{x}$ | $\hat{x} = \dfrac{1}{x}$ | $y = \beta_0 + \beta_1 \hat{x}$ |
| Hyperbolic: $y = \dfrac{x}{\beta_0 + \beta_1 x}$ | $\hat{y} = \dfrac{1}{y};\;\; \hat{x} = \dfrac{1}{x}$ | $\hat{y} = \beta_1 + \beta_0 \hat{x}$ |

## 3. Parameter Estimation

Significant level: $\alpha$

Confidence interval: $1-\alpha$

**Table 7:** *Frequently used z-value.*

| $\alpha$ | 0.005 | 0.01 | 0.025 | 0.05 | 0.1 |
|---|---|---|---|---|---|
| $z_\alpha$ | 2.58 | 2.33 | 1.96 | 1.65 | 1.28 |

**Table 8:** *Parameter estimation table.*

| Estimation | Given | Confidence interval |
|---|---|---|
| $\mu$ | $\sigma$ | • $\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$<br>• $\mu \gtrless \bar{X} \mp z_\alpha \frac{\sigma}{\sqrt{n}}$ |
| | $s$ | • $\bar{X} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$<br>• $\mu \gtrless \bar{X} \mp t_{\alpha,n-1} \frac{s}{\sqrt{n}}$ |
| $\sigma^2$ | $s^2$ | • $\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$<br>• $\sigma^2 > \frac{(n-1)s^2}{\chi^2_{\alpha,n-1}}$<br>• $\sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha,n-1}}$ |
| $p$ | $\hat{p}$ | • $\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$<br>• $p \gtrless \hat{p} \mp z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ |

## 4. Hypothesis Testing

**Table 9:** *Hypothesis testing table.*

<table>
<tr><td rowspan="4">Test<br>$\mu$</td><td>$T = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$<br>Given: $\sigma$</td><td>$H_0: \mu = \mu_0$<br>$H_1: \mu \neq \mu_0$<br>Reject, $|t| > z_{\alpha/2}$</td><td>$H_0: \mu \leq \mu_0$<br>$H_1: \mu > \mu_0$<br>Reject, $t > z_\alpha$</td><td>$H_0: \mu \geq \mu_0$<br>$H_1: \mu < \mu_0$<br>Reject, $t < -z_\alpha$</td></tr>
<tr><td>$p$-value</td><td>$p = 2\mathrm{P}(T > |t|)$</td><td>$p = \mathrm{P}(T > t)$</td><td>$p = \mathrm{P}(T < t)$</td></tr>
<tr><td>$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$<br>Given: $s$</td><td>$H_0: \mu = \mu_0$<br>$H_1: \mu \neq \mu_0$<br>Reject, $|t| > t_{\alpha/2,n-1}$</td><td>$H_0: \mu \leq \mu_0$<br>$H_1: \mu > \mu_0$<br>Reject, $t > t_{\alpha,n-1}$</td><td>$H_0: \mu \geq \mu_0$<br>$H_1: \mu < \mu_0$<br>Reject, $t < -t_{\alpha,n-1}$</td></tr>
<tr><td>$p$-value</td><td>$p = 2\mathrm{P}(T > |t|)$</td><td>$p = \mathrm{P}(T > t)$</td><td>$p = \mathrm{P}(T < t)$</td></tr>
<tr><td rowspan="2">Test<br>$\sigma^2$</td><td>$T = \frac{(n-1)s^2}{\sigma_0^2}$<br>Given: $s^2$</td><td>$H_0: \sigma^2 = \sigma_0^2$<br>$H_1: \sigma^2 \neq \sigma_0^2$<br>Accept $H_0$,<br>$\mathcal{X}^2_{1-\alpha/2,n-1} < t < \mathcal{X}^2_{\alpha/2,n-1}$</td><td>$H_0: \sigma^2 \leq \sigma_0^2$<br>$H_1: \sigma^2 > \sigma_0^2$<br>Reject $H_0$,<br>$t > \mathcal{X}^2_{\alpha,n-1}$</td><td>$H_0: \sigma^2 \geq \sigma_0^2$<br>$H_1: \sigma^2 < \sigma_0^2$<br>Reject $H_0$,<br>$t < \mathcal{X}^2_{1-\alpha,n-1}$</td></tr>
<tr><td>$p$-value</td><td>$p = 2\min\{\mathrm{P}(T < t), \mathrm{P}(T > t)\}$</td><td>$p = \mathrm{P}(T > t)$</td><td>$p = \mathrm{P}(T < t)$</td></tr>
<tr><td rowspan="2">Test<br>$p$</td><td>$T = \frac{p_0 - \hat{p}}{\sqrt{\hat{p}(1-\hat{p})/n}}$<br>Given: $\hat{p}$</td><td>$H_0: p = p_0$<br>$H_1: p \neq p_0$<br>Reject, $|t| > z_{\alpha/2}$</td><td>$H_0: p \leq p_0$<br>$H_1: p > p_0$<br>Reject, $t > z_\alpha$</td><td>$H_0: p \geq p_0$<br>$H_1: p < p_0$<br>Reject, $t < -z_\alpha$</td></tr>
<tr><td>$p$-value</td><td>$p = 2\mathrm{P}(T > |t|)$</td><td>$p = \mathrm{P}(T > t)$</td><td>$p = \mathrm{P}(T < t)$</td></tr>
<tr><td>Test<br>$\mu_x = \mu_y$</td><td>$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$<br>Given:<br>$\sigma_x^2, \sigma_y^2, n, m$</td><td colspan="3">Test for equality of means of two normal population:<br>$H_0: \mu_x = \mu_y \rightarrow \mu_x - \mu_y = 0$<br>$H_1: \mu_x \neq \mu_y$<br>Reject $H_0$, $|t| > z_{\alpha/2}$, $p$-value: $p = 2\mathrm{P}(T > |t|)$</td></tr>
<tr><td>Test<br>$\beta$</td><td>$T = \sqrt{\frac{(n-2)S_{XX}}{SS}}|B|$<br>Given data</td><td colspan="3">Test whether or not Y-data depends on X-data ($Y = \alpha + \beta X + \varepsilon$):<br>$H_0: \beta = 0$<br>$H_1: \beta \neq 0$<br>Reject $H_0$, $|t| > t_{\alpha/2,n-2}$, $p$-value: $p = 2\mathrm{P}(T_{n-2} > |t|)$</td></tr>
<tr><td>Test<br>$m$</td><td>$T = \frac{n_0 - 0.5n}{0.5\sqrt{n}}$<br>Given: $n_0$</td><td colspan="3">Test whether or not the median is equal to a given value:<br>$H_0: m = m_0$<br>$H_1: m \neq m_0$<br>Reject $H_0$, $\alpha > p$-value, $p = 2\min\{\mathrm{P}(Z < t), \mathrm{P}(Z > t)\}$<br>*(Using normal distribution)*</td></tr>
</table>