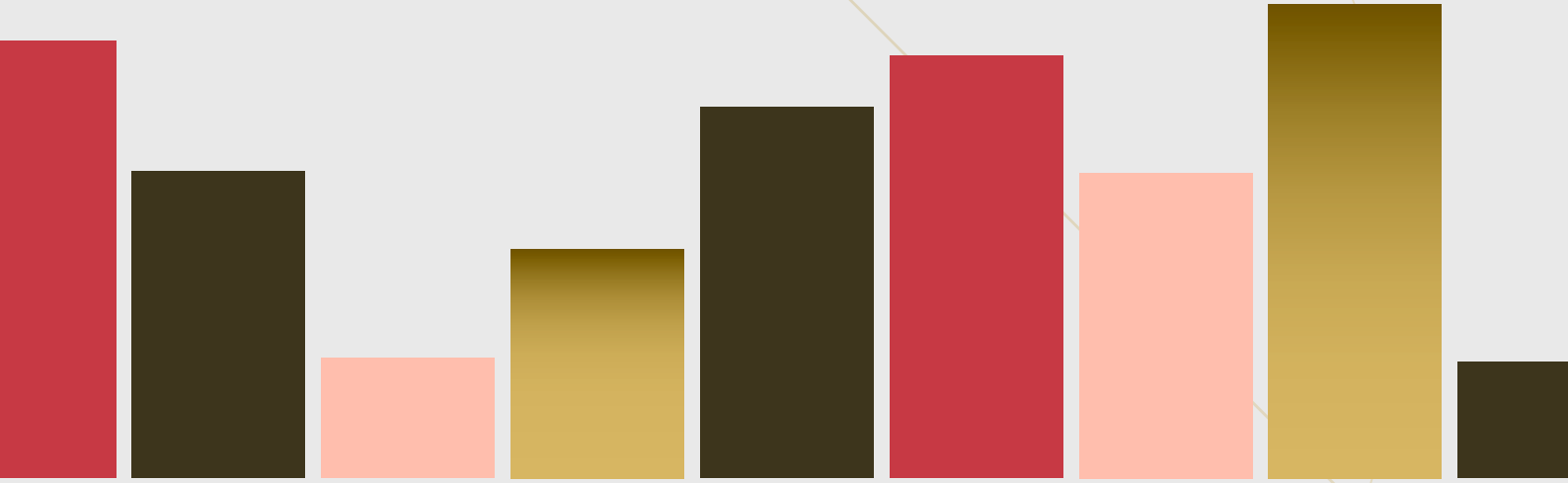
Chapter 3

Recap Sheet



• Definitions

- **Parameter (θ)**: a fixed (but unknown) number that describe the population (e.g., μ, σ^2, p, λ)
- **Statistic**: any function of the sample $X_1, \dots, X_n$ (e.g., $\overline{X}$, S_n^2, count of successes). A statistic is random because it depends on random data.
- **Point estimator T_n**: a statistic used to estimate a parameter with a single number $\hat{\theta} = T_n(X_1, \dots, X_n)$

Ex:
- $\hat{\mu} = \overline{X} = \frac{1}{n}\sum X_i$ (estimates μ)
- $\hat{p} = \overline{X}$ when $X_i \sim \text{Bernoulli}(p)$ (estimates p)
- $\hat{\lambda} = \frac{1}{\overline{X}}$ when $X_i \sim \text{Exp}(\lambda)$ (estimates λ)

• Properties of a good estimator

- **Unbiasedness**: T_n is an unbiased estimator of θ if
$$E[T_n] = \theta$$

Meaning: on average, across repeated samples, the estimator hits its true value.

If $E(T_n) - \theta \neq 0$, that difference is called bias:
- Positive bias → estimator tends to overestimate
- Negative bias → estimator tends to underestimate

- **Variance and MVUE**: The variance of the estimator tells us how concentrated its sampling distribution is around the true parameter.

Among all unbiased estimator, the one with the smallest variance is called the Minimum Variance Unbiased Estimator.

- **Mean Squared Error (MSE)**:

$$\begin{aligned} MSE(T_n) &= E[(T_n - \theta)^2] \\ &= \underbrace{Var(T_n)}_{\text{spread}} + \underbrace{(E(T_n) - \theta)^2}_{\text{bias}^2} \end{aligned}$$

• Methods of constructing point estimators

- **Method of moments**: Population moments depend on parameters, replace them with sample moments and solve for the parameter.

If $E(X) = \varphi(\theta)$, then the estimator is

$$\hat{\theta} = \varphi^{-1}(\overline{X})$$

Ex : We model values of sample by Binomial distrib. $B(2, p)$.
We calculated $\bar{x} = 2/3$. Let us find a point estimate for p.

For $B(2, p)$, $E(X) = np = 2p$. Using the method of moments :
$E(X) = \bar{X} \Leftrightarrow 2p = 2/3 \Leftrightarrow p = 1/3 = 0.333$

- **Maximum Likelihood Estimation** : Choose the parameter θ that makes the observed sample the most probable.

Likelihood function : $L(\theta; x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$

The MLE $\hat{\theta}$ is the value of θ that maximizes $L(\theta)$.

1. Likelihood ($L(\theta)$)
2. Log - likelihood ($l(\theta)$)
3. Derivative of log - likelihood ($l'(\theta)$)
4. Find maximum by solving $l'(\theta) = 0$

Ex : The number X of jumps is modeled by a Pascal (Geometrical) distribution on $\mathbb{N}^*$:

$$P(X = x) = p(1-p)^{x-1}, \quad x \geq 1$$

What is the MLE of p ?

For raw observations $x_1, \dots, x_n$, the likelihood is :

$$L(p) = \prod_{i=1}^{n} p(1-p)^{x_i - 1} = p^n (1-p)^{\sum_{i=1}^{n}(x_i - 1)} = p^n (1-p)^{\sum_{i=1}^{n}(x_i) - n}$$

We maximize the log - likelihood :

$$l(p) = \ln[L(p)] = \ln\left[p^n (1-p)^{\sum_{i=1}^{n}(x_i) - n}\right]$$

$$= \ln[p^n] + \ln\left[(1-p)^{\sum_{i=1}^{n}(x_i) - n}\right]$$

$$= n \ln(p) + \left(\sum_{i=1}^{n}(x_i) - n\right) \ln(1-p)$$

Derivative : $l'(p) = \frac{n}{p} - \frac{\sum_{i=1}^{n}(x_i) - n}{1-p}$

We solve :

$$l'(p) = 0 \Leftrightarrow \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - n}{1-p} = 0 \Leftrightarrow \frac{n}{p} = \frac{\sum_{i=1}^{n} x_i - n}{1-p}$$

$$\Leftrightarrow n(1-p) = p\left(\sum_{i=1}^{n} x_i - n\right)$$

$$\Leftrightarrow n - np = p\sum_{i=1}^{n} x_i - np$$

$$\Leftrightarrow n = p\sum_{i=1}^{n} x_i$$

as $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Therefore : $\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{n}{n \cdot \bar{x}} = \boxed{\frac{1}{\bar{x}}}$

• Usual Point Estimators

- **Normal Law $\mathcal{N}(\mu, \sigma^2)$**
  - Expected value : $E(x) = \mu$
  - Variance : $V(x) = \sigma^2$

→ Estimators : $\hat{\mu} = \bar{x} = \frac{1}{n} \sum x_i$ (unbiased)

Biased $\hat{\sigma}^2 = \tilde{S}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$

Unbiased $\hat{\sigma}^2 = \tilde{S}^{*2} = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{n-1}{n} \tilde{S}^2$

- **Binomial Law $B(n, p)$**
  - Expected value : $E(x) = np$
  - Variance : $V(x) = np(1-p)$

→ Estimator of p : $\hat{p} = \frac{\bar{x}}{n}$

- **Bernoulli Law $B(p)$**

Particular case of binomial with $n = 1$.

→ Estimator of p : $\hat{p} = \bar{x}$ (proportion of success individuals in sample)

- **Poisson Law $\mathcal{P}(\lambda)$**
  - $E(x) = V(x) = \lambda$

→ Estimator : $\hat{\lambda} = \bar{x}$

- **Exponential Law $\mathcal{E}(\lambda)$**
  - Expected value : $E(x) = 1/\lambda$
  - Variance : $V(x) = 1/\lambda^2$

→ Estimator : $\hat{\lambda} = 1/\bar{x}$

- **Uniform Law $U(0, \theta)$**
  - Expected value : $E(x) = \theta/2$
  - Variance : $V(x) = \theta^2/12$

→ Estimator (moments) : $\hat{\theta}_{mom} = 2\bar{x}$

→ Estimator (max likelihood) : $\hat{\theta}_{lik} = \max(x_i)$