Cheatsheet: Probability and Statistics

Seminar 2: Rules of Probability

Mutually Exclusive Events:

Events A and B are mutually exclusive (disjoint, incompatible) if:
$P(A \cap B) = 0$

Rules of Probability:

1. Complement Rule:
   $P(A^c) = 1 - P(A)$
2. Union Rule:
   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. Difference Rule:
   $P(A \setminus B) = P(A) - P(A \cap B)$

Conditional Probability:

The probability of A given B (if $P(B) \neq 0$):
$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

Independent Events:

Events A and B are independent if:
$P(A \cap B) = P(A)P(B) \quad \text{or} \quad P(A \mid B) = P(A)$

Total Probability Rule:

For a partition $\{A_i\}i \in I$ of the sample space S:
$P(E) = \sum i \in I P(A_i) P(E \mid A_i)$

Multiplication Rule:

For n events $A_1, A_2, \ldots, A_n$:
$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \ldots P(A_n \mid \bigcap_{i=1}^{n-1} A_i)$

Seminar 3: Probabilistic Models

1. Binomial Model:

• Describes the probability of k successes in n independent Bernoulli trials with success probability p

.
$P(n,k) = \binom{n}{k} p^k (1-p)^{n-k}$

2. Hypergeometric Model:

• Deals with the probability of drawing k successes from a finite population of N items, without replacement.

$P(n,k) = \frac{\binom{n_1}{k}\binom{N-n_1}{n-k}}{\binom{N}{n}}$

3. Poisson Model:

• Used for counting the probability of a given number of events occurring in fixed intervals of time/space.

• Formula involves the sum of probabilities for specific success occurrences in trials.

4. Pascal (Negative Binomial) Model:

• Describes the probability of achieving the n -th success after k failures in Bernoulli trials.

$P(n,k) = \binom{n+k-1}{n-1} p^n (1-p)^k$

5. Geometric Model:

• Represents the probability that the first success occurs after k failures.

$P(k) = p(1-p)^k$

Seminar 4: Discrete Random Variables

The expectation (expected value) of a discrete random variable X is calculated using the formula:

X	0	1	2	3
$P(X)$	$\frac{5}{18}$	$\frac{17}{36}$	$\frac{2}{9}$	$\frac{1}{36}$

$E(X) = \sum_i x_i \cdot P(X = x_i)$

Using the previously computed probabilities:

$E(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3)$

Substituting the values:

$E(X) = 0 \cdot \frac{5}{18} + 1 \cdot \frac{17}{36} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{1}{36}$

Seminar 5: Continuous Random Variables

1. Continuous Random Variables

A continuous random variable X has:

• Probability Density Function (PDF) $f(x)$:
$f(x) \geq 0$ for all x, and $\int_{-\infty}^{\infty} f(x)dx = 1.$

• Cumulative Distribution Function (CDF) $F(x)$:
$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$

• Probability for an interval:
$P(a < X < b) = \int_{a}^{b} f(x)dx.$

• $P(X = x) = 0$ for any specific value x.

2. Continuous Random Vectors

For (X, Y), a continuous random vector with joint PDF $f_{X,Y}(x, y)$ and joint CDF $F(x, y)$:
• Joint CDF:

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)dvdu.$$

• Marginal PDFs:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx.$$

• Independence:
X and Y are independent if:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

Transformation of Random Variables

To find the PDF of a transformed random variable $Y = g(X)$, where g is a differentiable and strictly monotone function, we use:
$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left|\frac{d}{dy}g^{-1}(y)\right|.$

or if g is not sitrctly monotone:
$f_Y(y) = \sum_i f_X\left(g_i^{-1}(y)\right) \cdot \left|\frac{d}{dy}g_i^{-1}(y)\right|.$

summing over all segments.

Here's the process:

1. Identify the function $Y = g(X)$.
2. Solve for X in terms of Y to get $X = g^{-1}(Y)$.
3. Compute $\frac{d}{dy} g^{-1}(y)$.
4. Substitute $g^{-1}(y)$ into $f_X(x)$ and multiply by $\left| \frac{d}{dy} g^{-1}(y) \right|$.

Seminar 7: Inequalities, Central Limit Theorem, Point Estimators

1. Markov's Inequality

For any non-negative random variable X and any $a > 0$:

$$P(|X| \geq a) \leq \frac{E(|X|)}{a}$$

2. Chebyshev's Inequality

For any random variable X with finite expectation $E(X)$ and variance $V(X)$, and for any $\varepsilon > 0$:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

3. Central Limit Theorem (CLT)

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables with mean $\mu = E(X_i)$ and standard deviation σ. Define the sum:

$$S_n = \sum_{i=1}^{n} X_i$$

As $n \to \infty$, the standardized sum:

$$Z_n = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

converges in distribution to the standard normal distribution:

$$Z_n \to N(0, 1)$$

This means the cumulative distribution function (CDF) of Z_n approaches the standard normal CDF $\Phi(z)$.

Steps for Solving Problems Using Methods of Moements (MoM)

Step 1: Find theoretical moments

• Compute the theoretical mean (first moment), variance (second moment), or higher moments from the given probability density function (PDF) or cumulative distribution function (CDF).

• Example:

$$E(X) = \int_{-\infty}^{\infty} x f(x;\theta)\, dx$$

Step 2: Compute sample moments

Calculate the sample moments from the given data:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Step 3: Solve for the parameter(s)

• Set the sample moments equal to the theoretical moments:

$$\text{Theoretical Moment} = \text{Sample Moment}$$

$$E(X) = \bar{X}$$

$$\int_{-\infty}^{\infty} x f(x;\theta)\, dx = \frac{1}{n}\sum_{i=1}^{n} X_i$$

• Solve the resulting equation for the unknown parameter(s).

Steps for Solving Problems Using MLE

Step 1: Write the likelihood function

• Given the PDF of the distribution, construct the likelihood function by multiplying the PDFs of the sample data:

$$L(\theta) = \prod_{i=1}^{n} f(X_i;\theta)$$

$$L(\theta) = f(X_1;\theta)\cdot f(X_2;\theta)\cdot \cdots \cdot f(X_n;\theta)$$

Step 2: Take the natural logarithm (log-likelihood function)

• The log-likelihood function simplifies calculations:

$\ln L(\theta) = \sum_{i=1}^{n} \ln f(X_i; \theta)$

Step 3: Differentiate the log-likelihood with respect to θ

• Compute the derivative of the log-likelihood with respect to the unknown parameter:

$\frac{d}{d\theta} \ln L(\theta) = 0$

Step 4: Solve for the parameter

• Solve the resulting equation to find the maximum likelihood estimator $\hat{\theta}$.

Standard Error of an Estimator:

The standard error of an estimator $\hat{\theta}$ is:

$\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$

Fisher Information:

$I_n(\theta) = -E\left[\frac{\partial^2 \ln L(X_1, ..., X_n; \theta)}{\partial \theta^2}\right]$

If the range of X does not depend on θ, then:

$I_n(\theta) = nI_1(\theta)$

Efficiency of an Estimator:

The efficiency of an estimator $\hat{\theta}$ is:

$e(\theta) = \frac{1}{I_n(\theta) V(\hat{\theta})}$

Estimator Properties:

An estimator $\hat{\theta}$ of the parameter θ is:

- Unbiased: if $E(\hat{\theta}) = \theta$.
- Absolutely Correct: if $E(\hat{\theta}) = \theta$ and $V(\hat{\theta}) \to 0$ as $n \to \infty$.

- MVUE (Minimum Variance Unbiased Estimator): if $E(\hat{\theta}) = \theta$ and its variance is the lowest among all unbiased estimators.
- Efficient: if $e(\theta) = 1$, meaning it achieves the Cramér-Rao lower bound.

Note: If an estimator is efficient, it is also the MVUE.