

## 1 Finite-sample properties

### Problem 1

Assume that  $X_i$  are i.i.d. with  $\mathbb{E}[X_i] = \mu$  and  $\text{var}[X_i] = \sigma^2$ . Define the sample average as  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  with  $i = 1, \dots, n$ .

- (a) Find  $\mathbb{E}[\bar{X}_n]$ .
- (b) Find  $\text{var}[\bar{X}_n]$ .
- (c) Show that

$$\mathbb{E} \left[ \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \right] = 0 \quad \text{and} \quad \text{var} \left[ \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \right] = 1.$$

### Problem 2

Assume that  $X_i \sim \mathcal{N}(\mu_x, \sigma_x^2)$  with  $i = 1, \dots, n_x$ ,  $Y_i \sim \mathcal{N}(\mu_y, \sigma_y^2)$  with  $i = 1, \dots, n_y$  and  $X_i \perp Y_i$ .

- (a) Find  $\mathbb{E}[\bar{X}_n - \bar{Y}_n]$ .
- (b) Find  $\text{var}[\bar{X}_n - \bar{Y}_n]$ .
- (c) Find the distribution of  $\bar{X}_n - \bar{Y}_n$ .
- (d) Find the distribution of  $\bar{X}_n + \bar{Y}_n$ .
- (e) Find the distribution of  $S_x = \frac{n_x \hat{\sigma}_x^2}{\sigma_x^2}$  where  $\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^{n_x} (X_i - \bar{X}_n)^2$  is the sample variance.
- (f) Find the distribution of  $\frac{n_x \hat{\sigma}_x^2}{(n_x - 1) \sigma_x^2} \bigg/ \frac{n_y \hat{\sigma}_y^2}{(n_y - 1) \sigma_y^2}$  using the following theorem:

**Theorem 1** (Ratio of  $\chi^2$  as F-distribution). Let  $Z_1 \sim \chi_m^2$ ,  $Z_2 \sim \chi_p^2$  where  $m, p \in \mathbb{N}$  and  $Z_1 \perp Z_2$ . Then

$$\frac{Z_1/m}{Z_2/p} \sim F(m, p),$$

that is, ratio has an F-distribution with  $m$  and  $p$  degrees of freedom.

Which kind of hypothesis could we test using the statistic above?

### Problem 3

Suppose that the random variables  $Y_1, \dots, Y_n$  satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

where  $x_1, \dots, x_n$  are fixed constants, and  $\epsilon_1, \dots, \epsilon_n$  are i.i.d.  $\mathcal{N}(0, \sigma^2)$  with  $\sigma^2$  unknown.

- (a) Show that  $\hat{\beta} = \sum Y_i / \sum x_i$  is an unbiased estimator of  $\beta$ .
- (b) Show that  $\tilde{\beta} = [\sum Y_i / x_i] / n$  is also an unbiased estimator of  $\beta$ .
- (c) Calculate the exact variances of estimators from (a) and (b). Which one would you prefer?

### Problem 4

Assume that  $X_1, \dots, X_n$  are i.i.d. with  $\mathcal{N}(\mu, \sigma^2)$ . Find a constant  $c$  that satisfies  $\mathbb{E}[g(S^2)] = \sigma$  where  $g(S^2) = c\sqrt{S^2}$  is the function of the sample variance.

## 2 Maximum likelihood estimation

### Problem 1

A random variable  $X$  is said to have a Pareto distribution with parameter  $\beta$ , denoted as  $X \sim \text{Pareto}(\beta)$ , if it is continuously distributed with density

$$f_X(x; \beta) = \begin{cases} \beta x^{-\beta-1}, & \text{if } x > 1, \\ 0, & \text{otherwise.} \end{cases}$$

A random sample  $x_1, \dots, x_N$  from the  $\text{Pareto}(\beta)$  population is available. Derive the maximum-likelihood estimator of  $\beta$ . Does it maximize the log-likelihood function?

### Problem 2

Suppose that the random variables  $Y_1, \dots, Y_n$  satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

where  $x_1, \dots, x_n$  are fixed constants, and  $\epsilon_1, \dots, \epsilon_n$  are i.i.d.  $\mathcal{N}(0, \sigma^2)$  with  $\sigma^2$  unknown. Find the MLE of  $\beta$ . Is it unbiased?

### Problem 3

Let  $x_1, \dots, x_N$  be a random sample from a gamma( $\alpha, \beta$ ) population.

- (a) Find the MLE of  $\beta$ , assuming  $\alpha$  is known.
- (b) If  $\alpha$  and  $\beta$  are both unknown, there is no explicit formula for the MLE of  $\alpha$ , but the maximum can be found numerically. How can we use the result in part (a) to reduce the problem to the maximization of a univariate function?

### Problem 4

Let  $x_1, \dots, x_N$  be a sample from the inverse Gaussian p.d.f.,

$$f_X(x; \mu, \lambda) = \left( \frac{\lambda}{2\pi x^3} \right)^{1/2} \exp \left\{ \frac{-\lambda(x - \mu)^2}{2\mu^2 x} \right\}, \quad x > 0.$$

Show that the MLEs of  $\mu$  and  $\lambda$  are

$$\hat{\mu}_{ML} = \bar{x}, \text{ and } \hat{\lambda}_{ML} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i} - \frac{1}{\bar{x}}}.$$

### Problem 5

The independent random variables  $X_1, \dots, X_N$  have the common distribution

$$P(X_i \leq x; \alpha, \beta) = \begin{cases} 0, & \text{if } x < 0, \\ \left(\frac{x}{\beta}\right)^\alpha, & \text{if } 0 \leq x \leq \beta, \\ 1, & \text{if } x > \beta. \end{cases}$$

where the parameters  $\alpha$  and  $\beta$  are positive. Find the MLEs of  $\alpha$  and  $\beta$ .

### Problem 6

Suppose that the random variables  $Y_1, \dots, Y_n$  satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

where  $x_1, \dots, x_n$  are fixed constants, and  $\epsilon_1, \dots, \epsilon_n$  are i.i.d.  $\mathcal{N}(0, \sigma^2)$  with  $\sigma^2$  unknown.

- Find the MLE of  $\sigma^2$ .
- Obtain the Fisher information matrix.

### Problem 7

Let  $X_1, \dots, X_N$  are identically and independently distributed as

$$f_X(x, \lambda) = \lambda \exp(-\lambda x).$$

- Find the MLE of  $\lambda$ .
- Obtain the Fisher information for  $\lambda$ . Does Information Equality Matrix holds in this case?
- Calculate the MLE of  $\lambda$  numerically.

## 3 Asymptotic theory

### Problem 1

Consider a random variable  $X_N$  with the probability distribution

$$X_N = \begin{cases} -N, & \text{with probability } \frac{1}{N}, \\ 0, & \text{with probability } 1 - \frac{2}{N}, \\ N, & \text{with probability } \frac{1}{N}. \end{cases}$$

- Does  $X_N \xrightarrow{p} 0$  as  $N \rightarrow +\infty$ ?
- Calculate  $\mathbb{E}[X_N]$  and  $\text{var}[X_N]$ . Now suppose that the distribution is

$$X_N = \begin{cases} 0, & \text{with probability } 1 - \frac{1}{N}, \\ N, & \text{with probability } \frac{1}{N}, \end{cases}$$

and calculate  $\mathbb{E}[X_N]$ .

- Is it true that if  $X_N \xrightarrow{p} 0$  as  $N \rightarrow +\infty$ , then  $\mathbb{E}[X_N] \rightarrow 0$ ?

## Problem 2

Assume having a random variable  $Y$  and a random sample  $y_1, \dots, y_N$ . Which statistics converge in probability by the Weak Law of Large Numbers (WLLN) and Continuous Mapping Theorem (CLT)? Existence of which moments do we need to assume?

a)  $\frac{1}{N} \sum_{i=1}^N y_i^2$ .

b)  $\frac{1}{N} \sum_{i=1}^N y_i^3$ .

c)  $\max_{1 \leq i \leq N} y_i$ .

d)  $\frac{1}{N} \sum_{i=1}^N y_i^2 - \left( \frac{1}{N} \sum_{i=1}^N y_i \right)^2$ .

e)  $\frac{\frac{1}{N} \sum_{i=1}^N y_i^2}{\frac{1}{N} \sum_{i=1}^N y_i}$ .

f)  $\mathbb{1}_{[0, +\infty)} \left( \frac{1}{N} \sum_{i=1}^N y_i \right)$ .

## Problem 3

Take a random sample  $x_1, \dots, x_N$ , where  $X > 0$ . Consider the sample geometric mean

$$\hat{\mu} = \left( \prod_{i=1}^N x_i \right)^{\frac{1}{N}}$$

and a population geometric mean

$$\mu = \exp(\mathbb{E}[\log(X)]).$$

Assuming  $\mu$  is finite, show that

$$\hat{\mu} \xrightarrow{p} \mu.$$

## Problem 4

Let  $X_1, X_2, \dots$  be a sequence of i.i.d.  $\mathcal{U}(0, 1)$  random variables. Define the sequence  $Y_N$  as

$$Y_N = \min(X_1, \dots, X_N).$$

Show that  $Y_N \xrightarrow{p} 0$  as  $N \rightarrow \infty$ .

## Problem 5

Suppose  $X_1, X_2, \dots$  are i.i.d. with mean and variance  $\mu$  and  $\sigma^2$ . Consider

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}_N)^2,$$

where  $\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i$ . Show that  $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$  as  $N \rightarrow \infty$ .

## 4 Generalized method of moments

### Problem 1

Let  $X_1, \dots, X_N$  be i.i.d. with finite fourth moment. Let  $\bar{X}_N$  and  $\bar{X}_N^2$  be the sample averages of  $X$  and  $X^2$  respectively. Find constants  $a$  and  $b$  and function  $c(N)$ , such that the vector sequence

$$c(N) \begin{pmatrix} \bar{X}_N - a \\ \bar{X}_N^2 - b \end{pmatrix}$$

converges to a nontrivial distribution, and determine this limiting distribution. Derive the asymptotic distribution of the sample variance  $\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}_N)^2$  using the Delta method.

### Problem 2

Denote by  $\hat{\alpha}_{MM}$  and  $\hat{\beta}_{MM}$  the method-of-moments (MM) estimators of  $\alpha$  and  $\beta$  parameters of the gamma distribution. Derive the asymptotic distribution of  $\hat{\theta}_{MM} = \begin{pmatrix} \hat{\alpha}_{MM} \\ \hat{\beta}_{MM} \end{pmatrix}$ . How can we obtain the asymptotic distribution using the Delta method?

### Problem 3

Let  $X_1, \dots, X_N$  be i.i.d. random variables with  $f_{X_i}(x; \lambda) = \lambda \exp(-\lambda x)$ .

- Recall the MLE for  $\lambda$ , derive its asymptotic distribution and the confidence interval for  $\hat{\lambda}_{ML}$ .
- Derive the MM estimator for  $\lambda$  and its asymptotic distribution.

## 5 Linear regression

### Problem 1

Let  $Y$  be a random variable that denotes the number of dots obtained when a fair six sided die is rolled. Let

$$X = \begin{cases} Y, & \text{if } Y \text{ is even,} \\ 0, & \text{if } Y \text{ is odd.} \end{cases}$$

- Find the joint distribution of  $\begin{pmatrix} X \\ Y \end{pmatrix}$ .
- Find the optimal predictor of  $Y$  given  $X$ .
- Find the optimal linear predictor.

### Problem 2

Answer the following questions as *true* or *false* and elaborate on your answer.

- Consider a sample  $\{x_i\}_{i=1}^N$  whose observations are all drawn from an *identical* random variable  $X$  with finite moments. For  $\mathbb{E}[\bar{X}] = \mathbb{E}[X]$  to hold, where  $\bar{X} = N^{-1} \sum_{i=1}^N X_i$  is the sample mean, this sample must be *random* (i.i.d.).
- Consider a univariate random (i.i.d.) sample  $\{X_1, \dots, X_N\}$ . The random variable  $X$  from which the sample is drawn has mean and variance denoted as  $\mathbb{E}[X]$  and  $\text{var}[X]$  respectively. Consider the following statistic  $S^*$ , which is a *variation* of the sample variance  $S^2$ ,

$$S^* = \frac{1}{N} \sum_{i=1}^N (X_i - \mathbb{E}[X])^2.$$

It turns out that  $\mathbb{E}[S^*] = \text{var}[X]$ .

- c. Consider two random (i.i.d.) samples  $\{x_i\}_{i=1}^{N_X}$  and  $\{y_i\}_{i=1}^{N_Y}$ . The first sample is drawn from a normally distributed random variable  $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ , while the second sample is drawn from another normally distributed random variable  $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ . To test that the scale parameters of the two distributions are equal ( $H_0 : \sigma_X^2 = \sigma_Y^2$ ) one can use the  $F$ -statistic:

$$F = \frac{S_X^2 / \sigma_X^2}{S_Y^2 / \sigma_Y^2}$$

and this is possible if the two samples have different size, i.e.  $N_X \neq N_Y$ .

- d. Consider a random (i.i.d.) sample  $\{x_i\}_{i=1}^N$  whose observations are drawn from some random variable  $X$ . Suppose that the Weak Law of Large Numbers applies:  $\bar{X} \xrightarrow{p} \mathbb{E}[X]$ . It follows that  $\bar{X} \xrightarrow{q.m.} \mathbb{E}[X]$ , i.e. the sample mean  $\bar{X}$  converges *in quadratic mean* to  $\mathbb{E}[X]$ .

### Problem 3

Consider a linear regression model with  $K$  regressors arranged in a vector  $x_i$ ,

$$\mathbb{E}[Y_i | x_i] = x_i' \beta_0 \Rightarrow \beta_0 = \mathbb{E}[x_i x_i']^{-1} \mathbb{E}[x_i Y_i].$$

Via the Continuous Mapping Theorem, prove the following asymptotic property about the average Total Sum of Squares (TSS),

$$\frac{1}{N} \sum_{i=1}^N e_i^2 = \frac{1}{N} e' e = \frac{1}{N} y' M y \xrightarrow{p} \mathbb{E}[Y_i^2] - \mathbb{E}[Y_i x_i' \beta_0].$$