Chapter 5

General Concepts of Parameter Estimation

Exercise 5.1: Consider a random variable x with expectation value μ and variance σ^2, and suppose we have a sample of n observations, $x_1, \ldots, x_n$. The purpose of this exercise is to show that the sample mean,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{5.1}$$

is a consistent estimator for the expectation value μ.

(a) The first step is to prove the Chebyshev inequality,

$$P(|x - \mu| \geq a) \;\leq\; \frac{\sigma^2}{a^2}, \tag{5.2}$$

which holds for any positive a as long as the variance of x exists. Do this by recalling the definition of the variance,

$$\sigma^2 \;=\; \int_{-\infty}^{\infty} (x - \mu)^2 \, f(x) \, dx, \tag{5.3}$$

where $f(x)$ is the p.d.f. of x. Use the fact that the integral (5.3) would be less if the region of integration were restricted to $|x - \mu| \geq a$, and would be even less if in that region, $(x - \mu)^2$ were to be replaced by a^2.

(b) Use the Chebyshev inequality to prove the weak law of large numbers, i.e. for any $\epsilon > 0$,

$$\lim_{n \to \infty} P\left(\left| \frac{1}{n}\sum_{i=1}^{n} x_i - \mu \right| \geq \epsilon \right) = 0. \tag{5.4}$$

This is equivalent to the statement that $\overline{x}$ is a consistent estimator for μ, and holds as long as the variance of x exists.

Exercise 5.2: Consider a random variable x of mean μ and variance σ^2, for which one has obtained sample of values $x_1, \ldots, x_n$.

(a) Suppose the mean μ has been estimated using the sample mean, $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Show that the sample variance,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{n}{n-1}\left(\overline{x^2} - \overline{x}^2\right), \tag{5.5}$$

is an unbiased estimator of the variance σ^2. (Use the fact that $E[x_i x_j] = \mu^2$ for $i \neq j$ and $E[x_i^2] = \mu^2 + \sigma^2$ for all i.)

(b) Suppose that the mean μ is known. Show that

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \overline{x^2} - \mu^2 \tag{5.6}$$

is an unbiased estimator for σ^2.

Exercise 5.3: (a) Show that the variance of s^2 (5.5) is

$$V[s^2] = E[s^4] - (E[s^2])^2 = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\mu_2^2\right), \tag{5.7}$$

where $\mu_k = E[(x-\mu)^k]$ is the kth central moment of x. To do this, first show that s^2 can be written as

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} x_i^2 - \frac{1}{n(n-1)}\sum_{i,j=1}^{n} x_i x_j. \tag{5.8}$$

Then show that the expectation value of s^4 is

$$E[s^4] = \frac{1}{(n-1)^2}\sum_{i,j=1}^{n} E[x_i^2 x_j^2] - \frac{2}{n(n-1)^2}\sum_{i,j,k=1}^{n} E[x_i x_j x_k^2] + \frac{1}{n^2(n-1)^2}\sum_{i,j,k,l=1}^{n} E[x_i x_j x_k x_l]. \tag{5.9}$$

Count how many terms in each sum give the algebraic moments μ'_4 or ${\mu'_2}^2$. Note that the rest of the terms all contain at least one power of μ. Express the result in terms of central moments μ_2 and μ_4 by setting the terms with μ equal to zero. Subtract the value of $(E[s^2])^2$ from Exercise (5.2) to obtain the final result.

(b) Find the variance of s^2 for the case where x follows a Gaussian distribution. Use the fact that the fourth central moment of a Gaussian is $\mu_4 = 3\sigma^4$.