Inequalities

Example 5.3. Upper Bounds in the Chebyshev Inequality. When X is known to take values in a range $[a, b]$, we claim that $\sigma^2 \leq (b-a)^2/4$. Thus, if σ^2 is unknown, we may use the bound $(b-a)^2/4$ in place of σ^2 in the Chebyshev inequality, and obtain

$$\mathbf{P}\big(|X-\mu| \geq c\big) \leq \frac{(b-a)^2}{4c^2}, \qquad \text{for all } c > 0.$$

To verify our claim, note that for any constant γ, we have

$$\mathbf{E}\big[(X-\gamma)^2\big] = \mathbf{E}[X^2] - 2\mathbf{E}[X]\gamma + \gamma^2$$

and the above quadratic is minimized when $\gamma = \mathbf{E}[X]$.

Problem 3.* Jensen inequality. A twice differentiable real-valued function f of a single variable is called **convex** if its second derivative $(d^2f/dx^2)(x)$ is nonnegative for all x in its domain of definition.

(a) Show that the functions $f(x) = e^{\alpha x}$, $f(x) = -\ln x$, and $f(x) = x^4$ are all convex.

(b) Show that if f is twice differentiable and convex, then the first order Taylor approximation of f is an underestimate of the function, that is,

$$f(a) + (x-a)\frac{df}{dx}(a) \leq f(x),$$

for every a and x.

(c) Show that if f has the property in part (b), and if X is a random variable, then

$$f\big(\mathbf{E}[X]\big) \leq \mathbf{E}\big[f(X)\big].$$

Central Limit Theorem

Problem 10. A factory produces X_n gadgets on day n, where the X_n are independent and identically distributed random variables, with mean 5 and variance 9.

(a) Find an approximation to the probability that the total number of gadgets produced in 100 days is less than 440.

(b) Find (approximately) the largest value of n such that

$$\mathbf{P}(X_1 + \cdots + X_n \geq 200 + 5n) \leq 0.05.$$

(c) Let N be the first day on which the total number of gadgets produced exceeds 1000. Calculate an approximation to the probability that $N \geq 220$.

Problem 11. Let $X_1, Y_1, X_2, Y_2, \ldots$ be independent random variables, uniformly distributed in the unit interval $[0, 1]$, and let

$$W = \frac{(X_1 + \cdots + X_{16}) - (Y_1 + \cdots + Y_{16})}{16}.$$

Find a numerical approximation to the quantity

$$\mathbf{P}\big(|W - \mathbf{E}[W]| < 0.001\big).$$

Classical Statistics

Example 9.3. Estimating the Parameter of an Exponential Random Variable. Customers arrive to a facility, with the ith customer arriving at time Y_i. We assume that the ith interarrival time, $X_i = Y_i - Y_{i-1}$ (with the convention $Y_0 = 0$) is exponentially distributed with unknown parameter θ, and that the random variables $X_1, \ldots, X_n$ are independent. (This is the Poisson arrivals model, studied in Chapter 6.) We wish to estimate the value of θ (interpreted as the arrival rate), on the basis of the observations $X_1, \ldots, X_n$.

Problem 4. Mixture models. Let the PDF of a random variable X be the mixture of m components:

$$f_X(x) = \sum_{j=1}^{m} p_j f_{Y_j}(x).$$

where

$$\sum_{j=1}^{m} p_j = 1, \qquad p_j \geq 0, \quad \text{for } j = 1, \ldots, m.$$

Thus, X can be viewed as being generated by a two-step process: first draw j randomly according to probabilities p_j, then draw randomly according to the distribution of Y_j. Assume that each Y_j is normal with mean μ_j and variance σ_j^2, and that we have a set of i.i.d. observations $X_1, \ldots, X_n$, each with PDF f_X.

(a) Write down the likelihood and log-likelihood functions.

(b) Consider the case $m = 2$ and $n = 1$, and assume that μ_1, μ_2, σ_1, and σ_2 are known. Find the ML estimates of p_1 and p_2.

(c) Consider the case $m = 2$ and $n = 1$, and assume that p_1, p_2, σ_1, and σ_2 are known. Find the ML estimates of μ_1 and μ_2.

(d) Consider the case $m \geq 2$ and general n, and assume that all parameters are unknown. Show that the likelihood function can be made arbitrarily large by choosing $\mu_1 = x_1$ and letting σ_1^2 decrease to zero. *Note:* This is an example where the ML approach is problematic.