**Tutorial 6 and 7 (Chapters 6 and 7)**

1. If 50% of the population of a large community is in favour of a proposed rise in school taxes, approximate the probability that a random sample of 100 people will contain at least 60 who are in favour of the proposition.

   **Solution**

   The number of people in favor of the proposition is $X \sim Bin(100, 0.5)$. The mean of $X$ is 50 and standard deviation is 5. So $P(X \geq 60) = P(\frac{X-50}{5} > \frac{60-50}{5}) \approx 1 - \Phi(2) \approx 0.023$. ( Since $P(X \geq 60) = P(X > 59)$, you can also plug in 59 in the approximation above.)

2. The lifetimes of interactive computer chips produced by a certain semiconductor manufacturer are normally distributed with parameters $\mu = 1.4 \times 10^6$ hours and $\sigma = 4 \times 10^5$. What is the approximate probability that a batch of 100 chips will contain at least 90 whose lifetimes are less than $1.8 \times 10^6$?

   **Solution** The probability that a computer chip whose lifetime is less than $1.8\times10^6$ is $\Phi(\frac{1.8\times10^6-1.4\times10^6}{4\times10^5}) = \Phi(1) \approx 0.84$.

   Denote by X the number of chips whose lifetimes are less than $1.8 \times 10^6$, then $X \sim Bin(100, 0.84)$. And $P(X < 90) \approx \Phi(\frac{90-84}{\sqrt{100\times0.84\times0.16}}) \approx 0.95$, so the answer is approximately 0.05.

3. (Optional) Suppose $X \sim Binomial(n, p)$. Use CLT to find the minimum value of $n$ that satisfies

$$P(|\frac{X}{n} - p| < \frac{\sqrt{Var(X)}}{2}) \geq 0.99$$

   **Solution**

   By CLT, $\frac{X/n-p}{\sqrt{Var(X/n)}} \sim N(0,1)$, thus $P(|\frac{X}{n} - p| < \frac{\sqrt{Var(X)}}{2}) = P(|\frac{X/n-p}{\sqrt{Var(X/n)}}| < \frac{\sqrt{Var(X)}}{2\sqrt{Var(X/n)}} \approx 2\Phi(\frac{\sqrt{Var(X)}}{2\sqrt{Var(X/n)}}) - 1 \geq 0.99$. From $\Phi(2.58) = 0.995$, we get $\frac{\sqrt{Var(X)}}{2\sqrt{Var(X/n)}} \geq 2.58$, which implies $n \geq 5.16$, so we should have $n \geq 6$.

4. Let $Y_1, \ldots, Y_n$ denote a random sample from the probability density function

$$f(y|\theta) = \begin{cases} (\theta+1)y^\theta, & 0 < y < 1; \theta > -1 \\ 0 & \text{otherwise} \end{cases}$$

   Find an estimator for $\theta$ by the method of moments and find the MLE.

   **Solution.** It is easy to show that $E[y] = \frac{\theta+1}{\theta+2}$ and thus the estimator is $\hat{\theta} = \frac{2\bar{Y}-1}{1-\bar{Y}}$. The likelihood is $L(\theta) = (\theta+1)^n(\prod_{i=1}^n Y_i)^\theta$ and the MLE is $\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln Y_i} - 1$.

5. If $Y_1, \ldots, Y_n$ denote a random sample from the normal distribution with known mean $\mu = 0$ and unknown variance $\sigma^2$, find the method-of-moments estimator of $\sigma^2$.

   **Solution.** Since $E[Y] = 0$ and $E[Y^2] = \sigma^2$, the estimator is $\frac{1}{n}\sum_{i=1}^n Y_i^2$.

6. If $Y_1, \ldots, Y_n$ denote a random sample from the normal distribution with mean $\mu$ and variance $\sigma^2$, find the method-of-moments estimators of $\mu$ and $\sigma^2$.

   **Solution.** Here, we have that $E[Y] = \mu$ and $E[Y^2] = \sigma^2 + \mu^2$. Thus $\hat{\mu} = \bar{Y}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n Y_i^2 - (\bar{Y})^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^2$.

7. Let $Y_1, \ldots, Y_n$ denote a random sample from the density function given by

$$f(y|\theta) = \begin{cases} \left(\frac{1}{\theta}\right) ry^{r-1}e^{-y^r/\theta}, & y > 0, \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $r$ is a known positive constant. Find the MLE of $\theta$.

**Solution.** The likelihood function is $L(\theta) = \theta^{-n} r^n (\prod_i Y_i)^{r-1} \exp\{-\sum_i Y_i^r/\theta\}$. The loglikelihood is $\ln L(\theta) = -n \ln \theta + n \ln r + (r-1)\ln(\prod_i Y_i) - \sum_i Y_i^r/\theta$. By taking a derivative w.r.t. $\theta$ and equating to 0, we find $\hat{\theta} = \frac{1}{n}\sum_i Y_i^r$.

8. Suppose that $Y_1, \ldots, Y_n$ constitute a random sample from a uniform distribution with probability density function

$$f(y|\theta) = \begin{cases} \frac{1}{2\theta+1} & 0 < y < 2\theta + 1 \\ 0 & \text{otherwise} \end{cases}$$

   (i) Obtain the MLE of $\theta$.

   (ii) Obtain the MLE for the variance of the this distribution.

   **Solution.** (i) The likelihood function is $L(\theta) = (2\theta+1)^{-n} I\{Y_{(n)} \le 2\theta+1\}$. The likelihood is maximized for small values of $\theta$. The smallest value that can maximize the likelihood is $\hat{\theta} = \frac{1}{2}(Y_{(n)} - 1)$, which is the MLE. (ii) Since $Var(Y) = \frac{(2\theta+1)^2}{12}$, by the invariance principle the MLE is $(Y_{(n)})^2/12$.

9. Suppose $Y_1, \ldots, Y_n$ is a random sample from the uniform distribution on $(\theta, \theta+1)$.

   (i) Show that $\bar{Y}$ is a biased estimator and compute the bias (the bias is defined as $E(\hat{\theta}) - \theta$).

   (ii) Find a function of $\bar{Y}$ that is an unbiased estimator of $\theta$.

   (iii) Find $MSE(\bar{Y})$ when $\bar{Y}$ is used as an estimator of $\theta$.

   **Solution** (i) $E[\bar{Y}] = \theta + 0.5$ and thus the bias is 0.5. (ii) $\bar{Y} - 0.5$ is an unbiased estimator. (iii) $Var(\bar{Y}) = \frac{1}{12n}$ and thus $MSE(\bar{Y}) = \frac{1}{12n} + 0.25$.

10. Suppose $Y \sim Bin(n, p)$. Then $Y/n$ is an unbiased estimator of $p$. To estimate the variance of $Y$, we can use $n(Y/n)(1 - Y/n)$.

   (i) Show that the suggested estimator is a biased estimator of $Var(Y)$.

   (ii) Modify $n(Y/n)(1 - Y/n)$ slightly to form an unbiased estimator of $Var(Y)$.

   **Solution** $E\{n(Y/n)(1 - Y/n)\} = EY - \frac{1}{n}E[Y^2] = np - p(1-p) - np^2 = (n-1)p(1-p)$. So the estimator $\frac{n}{n-1}n(Y/n)(1 - Y/n)$ is unbiased.

11. Let $Y_1, \ldots, Y_n$ be a random sample from a population with mean $\mu$ and variance $\sigma^2$. Consider the following estimators for $\mu$:

$$\hat{\mu}_1 = \frac{1}{2}(Y_1 + Y_2),\ \hat{\mu}_2 = \frac{1}{4}Y_1 + \frac{Y_2 + \cdots + Y_{n-1}}{2(n-2)} + \frac{1}{4}Y_n,\ \hat{\mu}_3 = \bar{Y}.$$

   (i) Show that all estimators defined above are unbiased.

   (ii) Find the variances of the estimators.

   **Solution** (i) $E(\hat{\mu}_1) = \frac{1}{2}(E[Y_1] + E[Y_2]) = \mu$; $E(\hat{\mu}_2) = \frac{\mu}{4} + \frac{(n-2)\mu}{2(n-2)} + \frac{\mu}{4}$; $E(\hat{\mu}_3) = E(\bar{Y}) = \mu$. (ii) $Var(\hat{\mu}_1) = \sigma^2/2$, $Var(\hat{\mu}_2) = \frac{\sigma^2}{8} + \frac{\sigma^2}{4(n-2)}$, $Var(\hat{\mu}_3) = \frac{\sigma^2}{n}$.