Problem Sheet 10

November 1, 2020

1. Let $X_1, \ldots, X_n$ be a random sample from a Bernoulli distribution with parameter θ. Find MLE of θ.

   Proof. The distribution of X_i is given by
$$p_X(x;\theta) = \begin{cases} \theta^x(1-\theta)^{1-x} & \text{for } x = 0, 1, \\ 0 & \text{otherwise.} \end{cases}$$
   Thus the likelihood function is given by
$$p(\theta) := p_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta) = p_{X_1}(x_1;\theta)\cdots p_{X_n}(x_n,\theta) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{\sum_{i=1}^n (1-x_i)}.$$
   Hence log likelihood function is given by
$$L(\theta) = \ln p(\theta) = \sum_{i=1}^n x_i \ln\theta + (n - \sum_{i=1}^n x_i)\ln(1-\theta).$$
   Setting $\frac{d}{d\theta}L(\theta) = 0$, we obtain $\theta = \frac{x_1+\cdots+x_n}{n}$.
   Hence MLE of θ is $\widehat{\Theta}_n = \frac{X_1+\cdots+X_n}{n} := \overline{X}$. □

2. Let $X_1, \ldots, X_n$ be a random sample from a Geometric distribution with parameter θ. Find MLE of θ.

   Proof. The log likelihood function is given by
$$L(\theta) = n\ln\theta + (\sum_{i=1}^n x_i - n)\ln(1-\theta).$$
   Setting $\frac{d}{d\theta}L(\theta) = 0$, we obtain MLE of θ is $\widehat{\Theta}_n = \frac{n}{X_1+\cdots+X_n} = \frac{1}{\overline{X}}$. □

3. Let $X_1, \ldots, X_n$ be a random sample from a Poisson distribution with parameter θ. Find MLE of θ.

Proof. The log likelihood function is given by

$$L(\theta) = -n\theta + \sum_{i=1}^{n} x_i \ln\theta + c,$$

here $c = -\ln(\prod_{i=1}^{n} x_i!)$ is independent of θ. Setting $\frac{d}{d\theta}L(\theta) = 0$, we obtain MLE of θ is $\widehat{\Theta}_n = \frac{X_1+\cdots+X_n}{n} = \overline{X}$. $\square$

4. Let $X_1, \ldots, X_n$ be a random sample from a Normal distribution with parameter $\mu = 1$ and σ^2. Find MLE of σ^2. Is the estimator unbiased?

   Proof. The log likelihood function is given by

   $$L(\sigma^2) = -\frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - 1)^2 + c,$$

   here $c = -\frac{n}{2}\ln(2\pi)$ is independent of σ^2. Setting $\frac{d}{d\sigma^2}L(\sigma^2) = 0$, we obtain MLE of σ^2 is $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1)^2$.
   Note $E[\widehat{\sigma^2}] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i - 1)^2] = E[(X-1)^2] = Var(X-1) + (E[X-1])^2 = Var(X) = \sigma^2$. Hence the estimator is unbiased. $\square$

5. Let $X_1, \ldots, X_n$ be a random sample from a Normal distribution with parameter μ and σ^2 (both μ and σ^2 unknown).

   Proof. The log likelihood function is given by

   $$L(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

   $$\frac{\partial}{\partial\mu}L(\mu, \sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu).$$

   $$\frac{\partial}{\partial\sigma^2}L(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

   (a) Find MLE of μ. Is it unbiased? Is it consistent? What the MSE of this estimator?
   We see MLE estimator for μ is $\widehat{\mu}_n = \frac{X_1+\cdots+X_n}{n} = \overline{X}$.
   Note, $E[\widehat{\mu}_n] = E[X] = \mu$, hence it is unbiased.
   Next $Var(\widehat{\mu}_n) = \frac{Var(X)}{n} = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$, hence it is consistent.
   Finally, $MSE = E[(\widehat{\mu}_n - \mu)^2] = Var(\widehat{\mu}_n) + (bias)^2 = \frac{\sigma^2}{n}$.

(b) Find MLE $\widehat{\sigma^2_{ML}}$ of σ^2. Show that $\widehat{\sigma^2_{ML}}$ is a biased estimator of σ^2. Show that $s^2 = \frac{n}{n-1}\widehat{\sigma^2_{ML}}$ is an unbiased estimator of σ^2. Show that $s = \sqrt{\frac{n}{n-1}\widehat{\sigma^2_{ML}}}$ is a biased estimator of σ. We call s^2 (resp. s) the sample variance (resp. sample standard deviation).

Setting both $\frac{\partial}{\partial \mu}L(\mu, \sigma^2) = \frac{\partial}{\partial \sigma^2}L(\mu, \sigma^2) = 0$, we see the MLE of σ^2 is given by

$$\widehat{\sigma^2_{ML}} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu}_n)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2.$$

Now $bias(\widehat{\sigma^2_{ML}}) = E[\widehat{\sigma^2_{ML}}] - \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} E[X_i^2] - E[\overline{X}^2] - \sigma^2 = E[X^2] - E[\overline{X}^2] - \sigma^2$. Note, $E[X^2] = Var(X) + (E[X])^2 = \sigma^2 + \mu^2$ and $E[\overline{X}^2] = Var(\overline{X}) + E[\overline{X}]^2 = \frac{\sigma^2}{n} + \mu^2$. Substituting everything together, we get $bias(\widehat{\sigma^2_{ML}}) = -\frac{1}{n}\sigma^2$.

Note $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Hence $E[s^2] = \frac{n}{n-1}E[\widehat{\sigma^2_{ML}}] = \sigma^2$. Thus s^2 is an unbiased estimator of σ^2.

Finally, note $0 < Var(s) = E[s^2] - (E[s])^2 = \sigma^2 - (E[s])^2$, hence $E[s] \neq \sigma$, hence s is a biased estimator of σ.

□

6. Let $X_1, \ldots, X_n$ be a random sample from the interval $[0, \theta]$ with $\theta > 0$, that is,

$$f_X(x; \theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \le x \le \theta, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $\widehat{\Theta}_n = \max\{X_1, \ldots, X_n\}$ is MLE of θ.

(b) Is $\widehat{\Theta}_n$ an unbiased estimator of θ?

(c) Is $\widehat{\Theta}_n$ consistent?

(d) Compute MSE of $\widehat{\Theta}_n$.

Proof. The likelihood function is given by

$$p(\theta) = \begin{cases} \frac{1}{\theta^n} & \text{if } 0 \le x_1, \ldots, x_n \le \theta, \text{ or equivalently } \theta \ge \max\{x_1, \ldots, x_n\}, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\frac{1}{\theta^n}$ is a decreasing function of θ, we see that $p(\theta)$ is maximum for $\theta = \max\{x_1, \ldots, x_n\}$.

Next we find the distribution of $Y = \widehat{\Theta}_n$.

$$F_Y(y) = P(\max\{X_1, \ldots, X_n\} \le y) = P(X_1 \le y) \cdots P(X_n \le y) = \begin{cases} 0 \text{ if } y < 0, \\ \left(\frac{y}{\theta}\right)^n \text{ if } 0 \le y \le \theta, \\ 1 \text{ if } y > 1. \end{cases}$$

Thus, $f_Y(y) = \begin{cases} \frac{n}{\theta^n} y^{n-1} \text{ if } 0 \le y \le \theta, \\ 1 \text{ otherwise.} \end{cases}$

We see $E[\widehat{\Theta}_n] = E[Y] = \int y f_Y(y) dy = \frac{n}{n+1}\theta$, hence the estimator is biased.
Note, $E[Y^2] = \frac{n}{n+2}\theta^2$, hence $Var(\widehat{\Theta}_n) = Var(Y) = \frac{n}{(n+1)^2(n+2)}\theta^2 \to 0$ as $n \to \infty$. Hence the estimator is consistent.
Finally, $MSE = Var(Y) + (E[Y] - \theta)^2 = \frac{2}{(n+1)(n+2)}\theta^2$. □

7. Let $X_1, \ldots, X_n$ be a random sample from the interval $(0, \theta)$ with $\theta > 0$, that is,

$$f_X(x; \theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < x < \theta, \\ 0 & \text{otherwise.} \end{cases}$$

Show that MLE of θ does not exist.

Proof. The likelihood function is given by

$$p(\theta) = \begin{cases} \frac{1}{\theta^n} & \text{if } \theta > \max\{x_1, \ldots, x_n\}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that this function does not achieve local maxima at any point, hence MLE does not exist. □

8. Let $X_1, \ldots, X_n$ be a random sample from the interval $[\theta, \theta + 1]$, that is,

$$f_X(x; \theta) = \begin{cases} 1 & \text{if } \theta \le x \le \theta + 1, \\ 0 & \text{otherwise.} \end{cases}$$

Show that MLE of θ is not unique.

Proof. The likelihood function is given by

$$p(\theta) = \begin{cases} 1 & \text{if } \theta \le x_1, \ldots, x_n \le \theta + 1, \text{ or equivalently } \max\{x_1, \ldots, x_n\} - 1 \le \theta \le \min\{x_1, \ldots, x_n\}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence any value in the interval $[\max\{X_1, \ldots, X_n\} - 1, \min\{X_1, \ldots, X_n\}]$ is a MLE of θ.

□

9. Let $X_1, \ldots, X_{1600}$ be random samples from a Normal distribution with mean μ and variance 25. Find the 95% confidence interval of μ.

Proof. Note $\Phi(1.96) = 0.975$, i.e. $z_{0.975} = 1.96$. Here $\sigma^2 = 25$, $n = 1600$. Thus the 95% confidence interval is given by

$$[\overline{X} - z_{0.975}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{0.975}\frac{\sigma}{\sqrt{n}}] = [\overline{X} - 0.245, \overline{X} + 0.245],$$

here $\overline{X} = \frac{X_1+\cdots+X_{1600}}{1600}$.

$\square$

10. Let $X_1, \ldots, X_n$ be random samples from a Normal distribution with mean μ and variance 25. How large must the sample size be so that the 95% confidence interval of μ has length 0.98?

Proof. The confidence interval is $[\overline{X} - z_{0.975}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{0.975}\frac{\sigma}{\sqrt{n}}]$, hence the length of the confidence interval is $2z_{0.975}\frac{\sigma}{\sqrt{n}}$, equating the last quantity with 0.98, we see that $n = 400$.

$\square$

11. Let $X_1, \ldots, X_{1600}$ be random samples from a distribution with mean μ and variance 25. Find approximate 95% confidence interval of μ.

Proof. Let $\overline{X} = \frac{X_1+\cdots+X_n}{n}$ be the estimator of μ. By CLT, limiting distribution of $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ is the distribution of standard normal random variable Z. Thus,

$$P(|\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}| < c) \approx P(|Z| < c).$$

For the confidence interval, we want c such that $P(|Z| < c) = 1-\alpha$, i.e., $\Phi(c) = 1-\frac{\alpha}{2}$, i.e. $c = z_{1-\frac{\alpha}{2}}$. Thus the approximate $100(1-\alpha)\%$ confidence interval of μ is given by

$$[\overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}].$$

Substituting $n = 1600, \sigma^2 = 25, \alpha = 0.05$, we see the required confidence interval is $[\overline{X} - 0.245, \overline{X} + 0.245]$. $\square$

12. Let $X_1, \ldots, X_{1600}$ be random sample from a Bernoulli distribution with parameter θ. Find approximate 95% confidence interval of θ.

Proof. <u>Method I:</u> We estimate mean θ by sample mean $\overline{X} = \frac{X_1+\cdots+X_n}{n}$. Using this, we estimate variance $\sigma^2 = \theta(1-\theta)$ by $\overline{X}(1-\overline{X})$. Note for large n, by CLT,

$$\frac{\overline{X}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \approx \frac{\overline{X}-\theta}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \approx Z = N(0,1).$$

Hence,

$$P\Big(\Big|\frac{\overline{X}-\theta}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}}\Big| < c\Big) \approx P(|Z| < c).$$

For the confidence interval, we want c such that $P(|Z| < c) = 1-\alpha$, i.e., $\Phi(c) = 1-\frac{\alpha}{2}$, i.e. $c = z_{1-\frac{\alpha}{2}}$. Thus the approximate $100(1-\alpha)\%$ confidence interval of μ is given by

$$[\overline{X} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\overline{X}(1-\overline{X})}, \overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\overline{X}(1-\overline{X})}].$$

Taking $\alpha = 0.05$ and $n = 1600$, we see the interval is given by $[\overline{X} - 0.049\sqrt{\overline{X}(1-\overline{X})}, \overline{X} + 0.049\sqrt{\overline{X}(1-\overline{X})}]$.

<u>Method II:</u> We estimate mean θ by sample mean $\overline{X} = \frac{X_1+\cdots+X_n}{n}$. Using this, we estimate variance σ^2 by sample variance $s_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$. Note for large n, by CLT,

$$\frac{\overline{X}-\theta}{\sigma/\sqrt{n}} \approx \frac{\overline{X}-\theta}{s_n/\sqrt{n}} \approx Z = N(0,1).$$

Hence,

$$P\Big(\Big|\frac{\overline{X}-\theta}{s_n/\sqrt{n}}\Big| < c\Big) \approx P(|Z| < c).$$

For the confidence interval, we want c such that $P(|Z| < c) = 1-\alpha$, i.e., $\Phi(c) = 1-\frac{\alpha}{2}$, i.e. $c = z_{1-\frac{\alpha}{2}}$. Thus the approximate $100(1-\alpha)\%$ confidence interval of μ is given by

$$[\overline{X} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}s_n, \overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}s_n].$$

Taking $\alpha = 0.05$ and $n = 1600$, we see the interval is given by $[\overline{X} - 0.049s_n, \overline{X} + 0.049s_n]$.

$\square$

Remark: Note that since X_i takes value 0 or 1, then $X_i^2 = X_i$, hence $\frac{n-1}{n}s_n^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^n X_i - \overline{X}^2 = \overline{X} - \overline{X}^2$. Since n is large, $\frac{n-1}{n}s_n^2 \approx s_n^2 = \overline{X}(1-\overline{X})$. Thus the intervals obtained using method I or method II are approximately same.

13. Let $X_1, \dots, X_{1600}$ be random sample from a distribution with mean μ and unknown variance $\sigma^2 < \infty$. If the sample mean is 76 and sample variance is 12, find approximate 95% confidence interval of μ.

Proof. We estimate mean θ by sample mean $\overline{X} = \frac{X_1 + \cdots + X_n}{n}$. Using this, we estimate variance σ^2 by sample variance $s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})^2$. Note for large n, by CLT,

$$\frac{\overline{X} - \theta}{\sigma/\sqrt{n}} \approx \frac{\overline{X} - \theta}{s_n/\sqrt{n}} \approx Z = N(0,1).$$

Hence,

$$P(\left|\frac{\overline{X} - \theta}{s_n/\sqrt{n}}\right| < c) \approx P(|Z| < c).$$

For the confidence interval, we want c such that $P(|Z| < c) = 1-\alpha$, i.e., $\Phi(c) = 1 - \frac{\alpha}{2}$, i.e. $c = z_{1-\frac{\alpha}{2}}$. Thus the approximate $100(1-\alpha)\%$ confidence interval of μ is given by

$$[\overline{X} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}} s_n, \overline{X} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}} s_n].$$

Taking $\alpha = 0.05$, $n = 1600$, $\overline{X} = 76$ and $s_n^2 = 12$ we see the interval is given by $[75.8303, 76.1697]$.

$\square$