MA20033 - Solution Sheet One

Simon Shaw
s.c.shaw@maths.bath.ac.uk

2004/05 Semester I

1. **State the sample space $\mathcal{S}$ and the parameter space Ω for the following random quantities. In each case, what is the interpretation of the parameters? Based on this, try to suggest an intuitive estimator for each parameter.**

   (a) $X \sim Bernoulli(p)$

   $\mathcal{S} = \{0, 1\}$, $\Omega = (0, 1)$. p is the probability of a success. Let $X_1, \ldots, X_n$ be iid $Bernoulli(p)$. An intuitive estimator might be the proportion of successes in the observations which is identical to the mean of the observations, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (b) $X \sim N(\mu, \sigma^2)$

   $\mathcal{S} = \Re$, $\Omega = \{\mu \in \Re, \sigma^2 \in \Re^+\}$. μ is the expected value and σ^2 the variance. Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$ then an intuitive estimator of μ is the mean of the X_i, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ while the variance of the observations, $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$, is an intuitive estimator of σ^2. Note that the sample variance, $\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, could equally be advocated. How might we compare these two plausible estimators?

   (c) $X \sim Bin(n, p)$

   $\mathcal{S} = \{0, 1, \ldots, n\}$, $\Omega = (0, 1)$. p is the probability of a success and n the number of trials. We assume n is known, and an intuitive estimator might be the observed proportion of successes, i.e. X/n. A (better) estimator may be to take $X_1, \ldots, X_m$ to be iid $Bin(n, p)$ and an intuitive estimator of p is then $\overline{X}/n$ where $\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$.

   (d) $X \sim \exp(\lambda)$

   $\mathcal{S} = \Re^+$, $\Omega = \Re^+$. I should have made clear the type of exponential I was considering. In the case that $f(x|\lambda) = \lambda \exp(-\lambda x)$ for $x \geq 0$ and 0 otherwise, then λ denotes the rate and an intuitive estimator might be $1/\overline{X}$ where $X_1, \ldots, X_n$ are iid $\exp(\lambda)$ and $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. If $f(x|\lambda) = \frac{1}{\lambda}\exp(-x/\lambda)$ for $x \geq 0$ and 0 otherwise then λ is the expected value and an intuitive estimator might be the sample mean, $\overline{X}$.

   (e) $X \sim Geo(p)$

   $\mathcal{S} = \{1, 2, \ldots\}$, $\Omega = (0, 1)$. p is the probability of a success, so an intuitive estimator might be the reciprocal of the number of trials until a success, i.e. $1/X$.

A (better) estimator may be to take $X_1, \dots, X_n$ to be iid $Geo(p)$ and an intuitive estimator of p is then $1/\overline{X}$ where $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

(f) $X \sim U(a, b)$

$\mathcal{S} = (a, b)$, $\Omega = \{a \in \Re, b \in \Re, \text{such that } a < b\}$. a and b are the minimum and maximum values which X can take. Let $X_1, \dots, X_n$ be iid $U(a, b)$. An intuitive estimator for a might be $\min\{X_1, \dots, X_n\}$ while $\max\{X_1, \dots, X_n\}$ may be an intuitive estimator for b.

2. **Suppose $X_1, X_2, \dots, X_n$ are iid $N(\mu, \sigma^2)$ random quantities. Using the properties of independent Normals and expectation and variance operators, derive the sampling distribution of $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i$.**

The sum of independent normal random quantities is also a normal random quantity (see the example in §5.2 of your MA10032 notes or the additional handout). Thus, $\hat{\mu}$ is a normal distribution and we just need to find the corresponding parameters: $E(\hat{\mu}|\mu, \sigma^2)$ and $Var(\hat{\mu}|\mu, \sigma^2)$. Expectation is a linear operator, so

$$E\left(\frac{1}{n}\sum_{i=1}^n X_i \middle| \mu, \sigma^2\right) = \frac{1}{n}\sum_{i=1}^n E(X_i|\mu, \sigma^2) = \frac{1}{n}\sum_{i=1}^n \mu = \mu.$$

As the covariance may be viewed as an inner product space,

$$\begin{aligned} Var\left(\frac{1}{n}\sum_{i=1}^n X_i \middle| \mu, \sigma^2\right) &= Cov\left(\frac{1}{n}\sum_{i=1}^n X_i, \frac{1}{n}\sum_{j=1}^n X_j \middle| \mu, \sigma^2\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n Cov(X_i, X_j|\mu, \sigma^2) \\ &= \frac{1}{n^2}\left\{\sum_{i=1}^n Var(X_i|\mu, \sigma^2) + \sum_{i=1}^n\sum_{j\neq i}^n Cov(X_i, X_j|\mu, \sigma^2)\right\}. \end{aligned}$$

As the X_i are independent, then $Cov(X_i, X_j|\mu, \sigma^2) = 0$ for $i \neq j$. Thus,

$$\begin{aligned} Var\left(\frac{1}{n}\sum_{i=1}^n X_i \middle| \mu, \sigma^2\right) &= \frac{1}{n^2}\sum_{i=1}^n Var(X_i|\mu, \sigma^2) \\ &= \frac{1}{n^2}\sum_{i=1}^n \sigma^2 = \sigma^2/n. \end{aligned}$$

Hence, $\hat{\mu} \sim N(\mu, \sigma^2/n)$.

3. (a) **Suppose that in an attempt to estimate the proportion of the population who are left-handed, 100 people are surveyed of whom 8 are left-handed. Assuming a Binomial model, evaluate the probability that $X = 8$ under the three values $p = 0.01$, $p = 0.1$ and $p = 0.5$. If these are**

the only possible values of p, so $\Omega = \{0.01, 0.1, 0.5\}$, what is the maximum likelihood estimate of p?

If $p = 0.01$ then

$$P(X = 8|p = 0.01) \;=\; \binom{100}{8}(0.01)^8(0.99)^{92} \;=\; 7.381693771 \times 10^{-6}.$$

Thus, $L(0.01) = 7.381693771 \times 10^{-6}$. If $p = 0.1$ then

$$P(X = 8|p = 0.1) \;=\; \binom{100}{8}(0.1)^8(0.9)^{92} \;=\; 0.114823026.$$

Hence, $L(0.1) = 0.114823026$. If $p = 0.5$ then

$$P(X = 8|p = 0.5) \;=\; \binom{100}{8}(0.5)^{100} \;=\; 1.467974647 \times 10^{-19},$$

whence, $L(0.5) = 1.467974647 \times 10^{-19}$. For $\Omega = \{0.01, 0.1, 0.5\}$, the maximum likelihood estimate is $\hat{p} = 0.1$ as this is the value of $p \in \Omega$ with the largest corresponding likelihood.

(b) **Now write down the probability that $X = 8$ under an arbitrary value of $p \in \Omega = (0, 1)$. What is the maximum likelihood estimate of p? Comment on this value.**

In this general setting,

$$L(p) \;=\; P(X = 8|p) \;=\; \binom{100}{8} p^8(1-p)^{92}.$$

Differentiating with respect to p gives

$$\begin{aligned} L'(p) &= \binom{100}{8}\{8(1-p) - 92p\}p^7(1-p)^{91} \\ &= (8 - 100p)p^7(1-p)^{91}. \end{aligned}$$

The maximum likelihood estimate is thus $\hat{p} = 8/100$. This corresponds to the observed proportion of successes, the intuitive estimate suggested in question 1(c).

4. **Let $x_1, x_2, \ldots, x_n$ be a random sample (so the X_i are independent) from an exponential distribution with probability density function**

$$f(x|\tau) \;=\; \frac{1}{\tau}\exp(-x/\tau) \quad 0 \leq x < \infty$$

and zero otherwise, where $\tau > 0$.

(a) **Show that if the model is correct it is sufficient to know the sample size n and the sample mean $\overline{x}$ to evaluate the likelihood function for any value of τ.**

The likelihood function is

$$\begin{aligned} L(\tau) &= \prod_{i=1}^{n} \frac{1}{\tau} \exp(-x_i/\tau) \\ &= \frac{1}{\tau^n} \exp(-\sum_{i=1}^{n} x_i/\tau) \\ &= \frac{1}{\tau^n} \exp(-n\overline{x}/\tau) \end{aligned}$$

for $\tau > 0$. Hence, it is sufficient to know the values n and $\overline{x}$ to compute $L(\tau)$.

(b) **Find the maximum likelihood estimator $\hat{\tau}$ of τ in terms of $\overline{X}$.**

The log-likelihood is

$$l(\tau) = -nlog(\tau) - n\frac{\overline{x}}{\tau}.$$

Thus, differentiating with respect to τ,

$$l'(\tau) = -\frac{n}{\tau} + n\frac{\overline{x}}{\tau^2},$$

so that $\overline{x}$ solves $l'(\tau) = 0$. We may readily check that $l''(\overline{x}) = -n/\overline{x}^2 < 0$ so that $\hat{\tau} = \overline{X}$ is the maximum likelihood estimator.

(c) **Show that $\hat{\tau}$ is an unbiased estimator of τ; that is, $E(\hat{\tau}|\tau) = \tau$ for all $\tau > 0$.**

With the exponential in the form given, $E(X|\tau) = \tau$ as

$$\begin{aligned} E(X|\tau) &= \int_0^\infty x\frac{1}{\tau} \exp(-x/\tau)\, dx \\ &= [-(x+\tau)\exp(-x/\tau)]_0^\infty \;=\; \tau. \end{aligned}$$

Thus, for each i, $E(X_i|\tau) = \tau$. So, using the linearity of expectation,

$$E(\hat{\tau}|\tau) = \frac{1}{n}\sum_{i=1}^{n} E(X_i|\tau) \;=\; \tau.$$