Estimation of Parameters

1 Maximum Likelihood Estimation

Let X be a random variable having pdf or pmf $f(x;\theta)$, where θ is an unknown parameter. The *likelihood function* of θ, corresponding to a sample of size n, is

$$L(\theta) = f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta).$$

The value of $\theta = T(x_1,\ldots,x_n)$ (say) for which $L(\theta)$ is maximum is the *maximum likelihood estimate* of θ. The corresponding statistic $\hat{\theta} = T(X_1,\ldots,X_n)$ is the *maximum likelihood estimator* (mle) of θ.

Where applicable, we maximise $L(\theta)$ by solving the equation $\dfrac{d}{d\theta}L(\theta) = 0$ for θ. As $L(\theta)$ is the product of $f(x_1;\theta),\ldots,f(x_n;\theta)$, it is more often convenient to maximise the *log-likelihood function* $\log L(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta)$ – which yields the same estimator $\hat{\theta}$, since $\log$ is an increasing function.

Example 1.1. Let $X \sim \mathcal{E}(\theta)$ (an exponential distribution with parameter θ). Then the pdf of X is $f(x;\theta) = \theta e^{-\theta x}$, $x \geq 0$, where $\theta > 0$. The likelihood function (assuming a sample of size n) will be

$$L(\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

and therefore the log-likelihood function will be

$$\log L(\theta) = n \log\theta - \theta \sum_{i=1}^{n} x_i.$$

Differentiating this with respect to θ and equating to zero, we have

$$\frac{n}{\theta} - \sum_{i=1}^{n} x_i = 0$$

which implies that

$$\theta = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}.$$

Thus, the MLE of θ is

$$\hat{\theta} = \frac{1}{\overline{X}}.$$

Example 1.2. Let X be a random variable with pdf $f(x;\theta) = (1+\theta)x^{\theta}$, $0 < x < 1$, where $\theta > 0$. The likelihood function will be

$$L(\theta) = (1+\theta)^n x_1^{\theta} \cdots x_n^{\theta}$$

and hence

$$\log L(\theta) = n \log(1+\theta) + \theta \sum_{i=1}^{n} \log x_i.$$

Equating its derivative with respect to θ to zero, we have

$$\frac{n}{1+\theta} + \sum_{i=1}^{n} \log x_i = 0$$

which implies that the MLE of θ is

$$\hat{\theta} = \frac{n}{\sum\limits_{i=1}^{n} \log \frac{1}{X_i}} - 1.$$

2 Confidence Intervals

Definition 2.1. If $(X_1, \ldots, X_n)$ is a random sample of size n, then the *sample mean* is defined as

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

and the *sample variance* is defined as

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2.$$

Theorem 2.2. *Suppose that $X \sim N(\mu, \sigma^2)$. Then the following hold:*

1. $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

2. $\dfrac{\overline{X} - \mu}{S/\sqrt{n-1}} \sim t_{n-1}$.

3. $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2_n$.

4. $\dfrac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$.

Confidence intervals for the mean and variance of a normal population, under different conditions, are as described in the table given below.

Parameter	**Condition**	**Interval**
μ	σ^2 known	$\left(\overline{x}-\frac{a\sigma}{\sqrt{n}}, \overline{x}+\frac{a\sigma}{\sqrt{n}}\right)$ where $P[-a<Z<a]=p,\ Z\sim N(0,1)$
	σ^2 unknown	$\left(\overline{x}-\frac{bS}{\sqrt{n-1}}, \overline{x}+\frac{bS}{\sqrt{n-1}}\right)$ where $P[-b<T<b]=p,\ T\sim t_{n-1}$
σ^2	μ known	$\left(\frac{1}{b}\sum_{i=1}^{n}(x_i-\mu)^2, \frac{1}{a}\sum_{i=1}^{n}(x_i-\mu)^2\right)$ where $P[Z<a]=P[Z>b]=\frac{1-p}{2},\ Z\sim\chi^2_n$
	μ unknown	$\left(\frac{ns^2}{b}, \frac{ns^2}{a}\right)$ where $P[Z<a]=P[Z>b]=\frac{1-p}{2},\ Z\sim\chi^2_{n-1}$