PROPERTIES OF POINT ESTIMATORS AND METHODS OF ESTIMATION

Maximum likelihood estimation 3:

Invariance and large-sample properties

$\mu, \sigma^2 \quad \nu$

σ^2

$\frac{\partial}{\partial \sigma}$

$\hat{\sigma}^2 = \frac{1}{n} \sum (Y_i - \bar{Y})^2$

$\hat{\sigma} = \sqrt{\frac{1}{n} \sum (Y_i - \bar{Y})^2}$

Invariance property of MLEs

- In some situations we might want to reparameterize, say $\nu = g(\theta)$ for some (known) function g.
- In such a case we could write the likelihood in terms of ν, take the log, find its derivative (with respect to ν), set it equal to zero, solve for ν, and check the 2nd derivative.
- One great property of maximum likelihood is that for any function g, the MLE of $g(\theta)$ is just $g(\hat{\theta})$!
- This is called the <u>invariance property</u> of MLEs

$\eta = \log \sigma$

$\hat{\eta} = \widehat{\log \sigma} = \log \hat{\sigma}$

$f(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$ $P(Y=0) = \boxed{e^{-\lambda}}$

Invariance property of MLE: Example

- Example 8: If $Y \sim \text{Poisson}(\lambda)$ then the probability of observing no events is $P(Y = 0) = e^{-\lambda}$. If $Y_1, Y_2, \ldots, Y_n \overset{iid}{\sim} Poisson(\lambda)$, find the MLE of this probability.

Find the MLE of λ

$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = e^{-n\lambda}\lambda^{\sum y_i}\left(\prod_{i=1}^{n}\frac{1}{y_i!}\right)$

$\ell(\lambda) = -n\lambda + \sum y_i \log\lambda + \log\left(\prod \frac{1}{y_i!}\right)$

$\frac{d}{d\lambda}\ell(\lambda) = -n + \frac{\sum y_i}{\lambda} + 0 \overset{set}{=} 0$

check 2^{nd} deriv. $\frac{d^2}{d\lambda^2}\ell(\lambda) = -\frac{\sum y_i}{\lambda^2} < 0$

$\frac{\sum y_i}{\lambda} = n$ $\lambda = \frac{\sum y_i}{n} = \bar{y}$

MLE of λ is $\hat{\lambda} = \bar{Y}$ MLE of $e^{-\lambda}$ $\boxed{e^{-\bar{Y}}}$

Large-sample properties of MLEs

- Theorem 3: If $Y_1, Y_2, \ldots, Y_n$ are iid with pdf $f(y|\theta)$, and if $\hat{\theta}$ is the MLE of θ, then (assuming some technical "regularity" conditions) $\hat{\theta}$ is consistent for θ. Furthermore, if n is large,

$$\hat{\theta} \sim N(\theta, \sigma^2_{\hat{\theta}}),$$

approx.

$$\lim_{n\to\infty} E[\hat{\theta}_n] = \theta$$

$$\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$$

where

$$\sigma^2_{\hat{\theta}} = \frac{1}{nE\left[-\frac{\partial^2 \log f(Y|\theta)}{\partial \theta^2}\right]}$$

- Note that this variance term depends on the unknown parameter θ. Typically, we just substitute $\hat{\theta}$ in place of θ.

(Large sample) confidence intervals with MLEs

standard error of $\hat{\theta}$

- Since $\hat{\theta} \dot{\sim} N(\theta, \sigma^2_{\hat{\theta}})$, an approximate $(1-\alpha)100\%$ confidence interval for θ is

$$\hat{\theta} \pm z_{\alpha/2}\boxed{\sigma_{\hat{\theta}}}.$$

- Example 9: If $Y_1, Y_2, \ldots, Y_n \sim Poisson(\lambda)$, find a 90% confidence interval for λ. Evaluate this interval if the observed data consist of 50 observations with mean 4.2.

iid

From before, we saw $\hat{\lambda} = \bar{Y}$

$$f(Y|\lambda) = e^{-\lambda}\frac{\lambda^{Y}}{Y!}$$

$$\log f(Y|\lambda) = -\lambda + Y\log\lambda - \log(Y!)$$

$$\frac{d}{d\lambda}\log f(Y|\lambda) = -1 + \frac{Y}{\lambda} \qquad \frac{d^2}{d\lambda^2}\log f(Y|\lambda) = -\frac{Y}{\lambda^2}$$

$$E\left[-\frac{\partial^2 \log f(Y|\lambda)}{\partial \lambda^2}\right] = E\left[\frac{Y}{\lambda^2}\right] = \frac{1}{\lambda^2} E[Y] = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

Poisson

$$\sigma_{\hat{\lambda}}^2 = \frac{1}{n\left(\frac{1}{\lambda}\right)} = \frac{\lambda}{n}$$

c.i. for λ $\quad \hat{\lambda} \pm z_{\alpha/2}\, \sigma_{\hat{\lambda}}$

replace with MLE $\bar{Y}$

$$\bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\lambda}{n}}$$

$$\boxed{\bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\bar{Y}}{n}}}$$

$n = 50$

$\bar{Y} = 4.2$

$$4.2 \pm 1.645 \sqrt{\frac{4.2}{50}}$$

$$4.2 \pm 0.48$$

$$\boxed{(3.72,\ 4.68)}$$