Chapter 3
Inference

(A) Likelihood

Let θ be a parameter to a density/mass $f(y;\theta)$ describing a model for the data

e.g., $\theta \in [0,1]$ in $Y_i \overset{iid}{\sim} \text{Bern}(\theta)$

$\longrightarrow$ $\theta = (\mu, \sigma^2) \in (\mathbb{R}, \mathbb{R}^+)$
for $Y_i \overset{iid}{\sim} N(\mu, \sigma^2)$

The <u>likelihood</u> is the joint probability of all data $y_1, \dots, y_n$, viewed as a function of the unknown parameter θ.

<u>Generically</u>: $L(\theta) = f(y_1, \dots, y_n ; \theta)$
$\overset{iid}{=} \prod_{i=1}^{n} f(y_i ; \theta)$

The value of θ that maximizes the likelihood (MLE)

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\ L(\theta) \equiv \underset{\theta}{\operatorname{argmax}}\ \underbrace{\log L(\theta)}_{l(\theta)}$$

For example: $Y_i \overset{iid}{\sim} N(\mu, \sigma^2) \quad i = 1, \ldots, n$

$$f(y_1, \ldots, y_n; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(y_i - \mu)^2\right\}$$

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right\}$$

$$l(\mu, \sigma^2) = c - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2$$

Now consider estimating μ via MLE

$$\frac{\partial l}{\partial \mu} = \frac{-2}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)(-1)$$

$$\phi \overset{set}{=} -n\hat{\mu} + \sum_{i=1}^{n} y_i \rightarrow \boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i}$$

$$= \bar{y}_n$$

Now consider estimating σ^2 via MLE

$$\frac{\partial l}{\partial \sigma^2} = \frac{-n}{2\sigma^2} - \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (y_i - \mu)^2$$

$$0 \overset{set}{=} -n - \frac{1}{\hat{\sigma}^2} \sum_{i=1}^{n} (y_i - \mu)^2$$

$$\boxed{\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \underset{\downarrow \bar{y}_n}{\hat{\mu}})^2}$$

One downside to this MLE is

$$\mathbb{E}\{\hat{\sigma}^2\} = \frac{n-1}{n} \sigma^2 \neq \sigma^2$$

(a biased estimate)

$$\boxed{s^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2}$$ is unbiased

generally preferred

Chapter 3
Continued

Ⓑ Inference

$\underline{Y_i \sim N(\mu, \sigma^2)}$ $\quad i = 1, \dots, n$

Estimator

$\underline{\hat{\mu}} = \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$

upper-case RV

$$\mathbb{E}\{\hat{\mu}\} = \mathbb{E}\{\bar{Y}\} = \mathbb{E}\{\tfrac{1}{n} \sum Y_i\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} E\{Y_i\} = \frac{1}{\cancel{n}} \cdot \cancel{n}\, \mu = \mu$$

$\boxed{\mathbb{E}\{\hat{\mu}\} = \mu}$ is unbiased

$$Var\{\hat{\mu}\} = Var\{\bar{Y}\} = Var\{\tfrac{1}{n} \sum Y_i\}$$

$$\overset{\text{indep}}{=} \frac{1}{n^2} \sum Var\{Y_i\} = \frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot \sigma^2 = \frac{\sigma^2}{n}$$

$\boxed{Var\{\hat{\mu}\} = \sigma^2/n}$

A sum of indep Gaussians is Gaussian.

$$\hat{\mu} = \bar{Y} \sim N(\mu, \sigma^2/n)$$

Assuming that σ^2 is known, we can compare this <u>sampling distribution</u> to an estimate from data under $H_0: \mu = \mu_0$.

<u>Standardize</u>

$$z = \frac{\bar{y} - \mu}{\sigma/\sqrt{n}} \qquad z \sim N(0,1)$$

$$\frac{\text{est} - \text{wtk}}{\text{se}} \quad (\mu_0)$$

It is unsatisfying to assume a particular value of σ^2, when we have a good estimate s^2.

However, incorporating uncertainty about s^2 involves understanding its distribution.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

If $Y_i \sim N(0,1)$ then $Y_i^2 \sim \chi_1^2$

$Y_i \sim N(\mu, \sigma^2)$ $\qquad \dfrac{(Y_i-\mu)^2}{\sigma^2} \sim \chi_1^2$

so $\bar{Y} \sim N(\mu, \sigma^2/n)$ $\qquad \dfrac{\bar{Y}-\mu}{\sigma^2/n} \sim \chi_1^2$

If $Y_1, \ldots, Y_n \overset{iid}{\sim} N(0,1)$ then $\sum_{i=1}^n Y_i^2 \sim \chi_n^2$

$Y_1, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$ $\qquad \dfrac{\Sigma(Y_i-\mu)^2}{\sigma^2} \sim \chi_n^2$

A homework exercise asks you to show:

$$\sum_{i=1}^n \frac{(Y_i-\mu)^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2} + \frac{(\bar{Y}-\mu)^2}{\sigma^2/n}$$

$$\sim \chi_n^2 \qquad (\sim \chi_{n-1}^2) \qquad \sim \chi_1^2$$

Therefore $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$

so $\boxed{\sigma^2 \sim (n-1)s^2 / \chi_{n-1}^2}$

An English statistician working at Guinness named Gosset, publishing under the pseudonym "Student", showed that "convolving" a Gaussian variance over a $(\chi^2)^{-1}$ distribution

$$Y \sim N(\mu, \sigma^2) \qquad \sigma^2 \sim (n-1)s^2 / \chi^2_{n-1}$$

has a closed form

$\rightarrow$ student-t distribution.

$$t = \frac{\bar{y} - \mu}{s/\sqrt{n}} \qquad T \sim t_{n-1}$$

$$\frac{\text{est - wtk}}{\text{se}}$$

Chapter 3
Continued

© Confidence Intervals

A <u>confidence interval</u> is the inverse of a hypothesis test.

- easy by simulation
- swap the role of μ_0 & $\hat{\mu}$

<u>Analytically?</u> (95% CI)

Let $q_{\alpha/2}$ & $q_{1-\alpha/2}$ denote the central 95% quantiles of a student-t distribution with $r=n-1$ dof.

a good approx is $q_{\alpha/2} = -2$

$q_{1-\alpha/2} = +2$

Consider $T \equiv \dfrac{\bar{Y} - \mu}{s/\sqrt{n}} \sim t_{n-1}$

$$1-\alpha = \mathbb{P}\left(q_{\alpha/2} \leq T \leq q_{1-\alpha/2} \right)$$

How to "invert" that so that μ is in the middle of some bounds.

$$1-\alpha = \mathbb{P}\left(q_{\alpha/2} \leq \frac{\bar{Y} - \mu}{s/\sqrt{n}} \leq q_{1-\alpha/2} \right)$$

$$= \mathbb{P}\left(\bar{Y} + \frac{s}{\sqrt{n}} q_{\alpha/2} \leq \mu \leq \bar{Y} + \frac{s}{\sqrt{n}} q_{1-\alpha/2} \right)$$

$$\bar{y} \pm q_{1-\alpha/2}\, s/\sqrt{n} \equiv CI$$