

# STAT 3100 Lecture Note

Week Twelve (Nov 29 & Dec 1, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Dec/01 20:42:54

---

## Asymptotic properties of MLE (con'd)

### Asymptotic efficiency of MLE (CB Thm 10.1.12 & Ex. 10.7)

- $\sqrt{n}(\hat{\theta}_{\text{ML}} - \theta_0) \xrightarrow{d} \mathcal{N}(0, 1/I_1(\theta_0))$ , provided that  $\hat{\theta}_{\text{ML}}$  is the MLE for  $\theta_0$ , we have the previous four regularity conditions (for the consistency of MLE) plus the following two more (CB Sec 10.6.2):
  - For each  $x \in \text{supp}(X)$ ,  $f(x | \theta)$  is three times continuously differentiable with respect to  $\theta$ ; and  $\int f(x | \theta) dx$  can be differentiated three times under the integral sign;
  - for each  $\theta \in \Theta$ , there exists  $c(\theta) > 0$  and  $M(x, \theta)$  such that  $|\frac{\partial^3}{\partial \theta^3} \ln f_X(x | \theta)| \leq M(x, \theta)$  for all  $x \in \text{supp}(X)$  and  $\theta \in (\theta - c(\theta), \theta + c(\theta))$ .
- In practice,
  - $nI_1(\theta_0) = I_n(\theta_0) \approx I_n(\hat{\theta}_{\text{ML}}) \approx \hat{I}_n(\hat{\theta}_{\text{ML}})$ 
    - \* (Expected) Fisher information (number)  $I_n(\theta_0) = -E\{H(\theta_0; \mathbf{X})\}$
    - \* Observed Fisher information (number)  $\hat{I}_n(\hat{\theta}_{\text{ML}}) = -\frac{\partial^2}{\partial \theta^2} \ln L(\theta; \mathbf{x})|_{\theta=\hat{\theta}_{\text{ML}}} = -H(\hat{\theta}_{\text{ML}}; \mathbf{x})$
  - Hence  $\text{var}(\hat{\theta}_{\text{ML}}) \approx 1/I_n(\theta_0) \approx 1/I_n(\hat{\theta}_{\text{ML}}) \approx 1/\hat{I}_n(\hat{\theta}_{\text{ML}})$

### Delta method

- (CB Thm 5.5.24, delta method) If  $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ ,  $\tau$  is NOT a function of  $n$ , and  $\tau'(\theta) \neq 0$ , then
$$\sqrt{n}\{\tau(\hat{\theta}_n) - \tau(\theta)\} \xrightarrow{d} \mathcal{N}(0, \{\tau'(\theta)\}^2 \sigma^2).$$
  - Hence  $\text{var}\{\tau(\hat{\theta}_n)\} \approx \{\tau'(\hat{\theta}_n)\}^2 \sigma^2 / n$  if  $\tau'(\theta) \neq 0$
- (CB Thm 5.5.26, second-order delta method) If  $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ ,  $\tau$  is NOT a function of  $n$ ,  $\tau'(\theta) = 0$ , and  $\tau''(\theta) \neq 0$ , then

$$n\{\tau(\hat{\theta}_n) - \tau(\theta)\} \xrightarrow{d} \frac{\tau''(\theta)\sigma^2}{2} \chi^2(1).$$

- Hence  $\text{var}\{\tau(\hat{\theta}_n)\} \approx \{\tau''(\hat{\theta}_n)\}^2 \sigma^4 / (2n^2)$  if  $\tau'(\theta) = 0$  but  $\tau''(\theta) \neq 0$

### CB Example 10.1.17 & Ex. 10.9

- iid  $X_1, \dots, X_n \sim p(x | \lambda) = \lambda^x \exp(-\lambda) / x!$ ,  $x \in \mathbb{Z}^+$ ,  $\lambda > 0$ . To estimate  $\Pr(X_i = 0) = \exp(-\lambda)$ .
  - a. Consider  $T_n = n^{-1} \sum_i \mathbf{1}_{\{0\}}(X_i)$  and MLE  $W_n = \exp(-\bar{X}_n)$ . Compute  $\text{ARE}(T_n, W_n)$ , the ARE of  $T_n$  with respect to  $W_n$ .
  - b. Find the UMVUE for  $\Pr(X_i = 0)$ , say  $U_n$ , and then calculate  $\text{ARE}(U_n, W_n)$ .
    - Hint:  $\sqrt{n}(U_n - W_n) \xrightarrow{P} 0$  (derived from S. Portnoy, *The Annals of Statistics*, 1977, 5, pp. 522–529, Theorem 1) and  $\sum_{i=1}^n X_i \sim \text{Poisson}(n\lambda)$

## Approximation to the variance of $\hat{\theta}_n$

- Why?
  - Reflect the variation or dispersion of  $\hat{\theta}_n$
  - Help approximate the distribution of  $\hat{\theta}_n$  (and further construct the confidence region for  $\theta$ ) if assuming normality
- How?
  - Utilizing the asymptotic variance of  $\hat{\theta}_n$
  - Resampling methods, e.g., bootstrapping

### CB Example 10.1.17 & Ex. 10.9 (con'd)

- iid  $X_1, \dots, X_n \sim p(x | \lambda) = \lambda^x \exp(-\lambda)/x!$ ,  $x \in \mathbb{Z}^+$ ,  $\lambda > 0$ . Define  $\theta = \Pr(X_i = 2 | \lambda) = \lambda^2 \exp(-\lambda)/2$ . Approximate the variance of  $\hat{\theta}_{\text{ML}} = \bar{X}_n^2 \exp(-\bar{X}_n)/2$  by delta methods.

### CB Example 10.1.15

- Holding iid  $X_i \sim \text{Bernoulli}(p)$ , the variance of  $\text{Bernoulli}(p)$  is  $\tau(p) = p(1-p)$  whose MLE is  $\tau(\hat{p}_{\text{ML}}) = \bar{X}_n(1 - \bar{X}_n)$ . Approximate  $\text{var}\{\tau(\hat{p}_{\text{ML}})\}$  by delta methods.

### Bootstrapping the variance of $\hat{\theta}_n$ (CB Sec. 10.1.4)

- Nonparametric bootstrap:
  1. For  $j$  in  $1 : B$ , do steps 2–3.
  2. Draw the  $j$ th resample  $\mathbf{x}_j^*$  of size  $n$  from the original sample  $\mathbf{x} = \{x_1, \dots, x_n\}$ , with replacement, i.e., create a new iid sample  $\mathbf{x}_j^*$  from  $F_n$  (the empirical cdf of the original sample)
  3. Let  $\hat{\theta}_j^* = \hat{\theta}(\mathbf{x}_j^*)$ .
  4.  $\text{var}(\hat{\theta}) \approx$  the sample variance of  $\{\hat{\theta}_1^*, \dots, \hat{\theta}_B^*\}$ .
- Parametric bootstrap:
  1. For  $j$  in  $1 : B$ , do steps 2–3.
  2. Draw the  $j$ th resample  $\mathbf{x}_j^*$  of size  $n$  from a fitted model  $f(x | \hat{\theta})$ .
  3. Let  $\hat{\theta}_j^* = \hat{\theta}(\mathbf{x}_j^*)$ .
  4.  $\text{var}(\hat{\theta}) \approx$  the sample variance of  $\{\hat{\theta}_1^*, \dots, \hat{\theta}_B^*\}$ .

### CB Example 10.1.15

- Holding iid  $X_i \sim \text{Bernoulli}(p)$ , the variance of  $\text{Bernoulli}(p)$  is  $\tau(p) = p(1-p)$  for which the MLE is  $\tau(\hat{p}_{\text{ML}}) = \bar{X}_n(1 - \bar{X}_n)$ . Approximate  $\text{var}\{\tau(\hat{p}_{\text{ML}})\}$  by the bootstrap.

```
options(digits = 4)
set.seed(1)
B = 1e4L
n = 30
x = rbinom(n, 1, prob = .7)
theta_ml = mean(x)
tau_theta_star_np = numeric(B)
tau_theta_star_p = numeric(B)
# Nonparametric bootstrap
for (j in 1:B){
  x_star = sample(x, size = n, replace = T)
  tau_theta_star_np[j] = mean(x_star)*(1-mean(x_star))
}
```

```

var(tau_theta_star_np)
# Parametric bootstrap
for (j in 1:B){
  x_star = rbinom(n, size = 1, prob = theta_ml)
  tau_theta_star_p[j] = mean(x_star)*(1-mean(x_star))
}
var(tau_theta_star_p)
# Estimate via the first-order delta method
theta_ml*(1-theta_ml)*(1-2*theta_ml)^2/n
# Estimate via the second-order delta method
2*theta_ml^2*(1-theta_ml)^2/n^2

```

## Large-sample hypothesis testing

### Recall the LRT

- $H_0 : \theta \in \Theta_0$  v.s.  $H_1 : \theta \in \Theta_1$ , where  $\Theta = \Theta_0 \cup \Theta_1$
- LRT statistic

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_0} L(\theta; x)}{\sup_{\theta \in \Theta} L(\theta; x)} = \frac{L(\hat{\theta}_{0,ML}; x)}{L(\hat{\theta}_{ML}; x)}$$

- $\hat{\theta}_{0,ML}$ : constrained MLE for  $\theta \in \Theta_0$
- $\hat{\theta}_{ML}$ : unconstrained MLE for  $\theta \in \Theta$
- $\{x : \lambda(x) \leq c_\alpha\}$ : rejection region of level  $\alpha$  LRT
  - $c_\alpha$  is such defined that  $\sup_{\theta \in \Theta_0} \Pr(\lambda(\mathbf{X}) \leq c_\alpha \mid \theta) = \alpha$

### Asymptotic LRT rejection region (CB Thm 10.3.1 & 10.3.3)

- Under  $H_0$ , as  $n \rightarrow \infty$ ,

$$-2 \ln \lambda(\mathbf{X}) \xrightarrow{d} \chi^2(\nu),$$

where  $\nu$  = difference of numbers of free parameters in  $\Theta_0$  and  $\Theta$ .

- (CB Thm 10.3.3)  $\{x : -2 \ln \lambda(x) \geq \chi_{\nu, 1-\alpha}^2\}$ : asymptotic rejection region of level  $\alpha$  LRT
  - $\chi_{\nu, 1-\alpha}^2$  is the  $1 - \alpha$  quantile of  $\chi^2(\nu)$ .

### CB Example 10.3.4

- iid  $X_1, \dots, X_n \sim f(x \mid p_1, \dots, p_5) = p_x$ ,  $x = 1, \dots, 5$ ,  $\sum_{k=1}^5 p_k = 1$  and  $p_k \in (0, 1)$ . i.e., the categorical distribution. Specify the level  $\alpha$  LRT rejection region for  $H_0 : p_1 = p_2 = p_3$  and  $p_4 = p_5$  vs.  $H_1$  : Otherwise.

```

options(digits = 4)
set.seed(1)
B = 1e4L
n = 1e3L
p0 = c(.1, .1, .1, .35, .35)
ys = rmultinom(B, n, p0)
ms = matrix(
  c(rep(colMeans(ys[1:3,]), times = 3), rep(colMeans(ys[4:5,]), times = 2)),
  nrow = length(p0), ncol = B, byrow = T)
test_stats = 2*colSums(ys*log(ys/ms))
seg = seq(0, 20, length.out=1000)

```

```
pdfchi2 = dchisq(seg, 3)
hist(test_stats, breaks=100, xlim=c(0,20),
      freq=F, xlab = expression(paste('-2ln', lambda, '(x)')), main = '')
lines(seg, pdfchi2, col = "red")
```

### Take-home exercises (NOT to be submitted; to be potentially covered in labs)

- CB Ex. 10.17(a–c), 10.36, 10.38
- HMC Ex. 6.3.16–6.3.18