STA 114: Statistics

Lab 3

This lab work is intended to illustrate normal approximation theory. You have seen a first instance of this theory in your probability course, in the form of the central limit theorem (CLT) that states that for IID random variables $X_1, X_2, \cdots, X_n$ with mean $\mu = \mathrm{E}X_i$ and variance $\sigma^2 = \mathrm{Var}X_i$, the sample mean $\bar{X}$ is approximately distributed as $\mathsf{Normal}(\mu, \sigma^2/n)$. More precisely, the random variable $Z_n = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ converges in distribution/law to $\mathsf{Normal}(0,1)$ as $n \to \infty$. i.e., $\lim_{n\to\infty} P(Z_n \le z) = \Phi(z)$ for every real number z.

For many regular statistical models, with differentiable log-likelihood functions that admit a unique maxima, the MLE often exhibits approximate normality. This is what we'll explore today. In particular, for a statistical model $X \sim f(x|\theta)$, $\theta \in \Theta$, and a fixed $\theta_0 \in \Theta$ we will see whether the distribution of $Z = \sqrt{I_X}(\hat{\theta}_{\text{MLE}}(X) - \theta_0)$ can be approximated by $\mathsf{Normal}(0,1)$, when $\theta = \theta_0$ is the truth.

Asymptotic Normality of the MLE

Consider the model $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Poisson}(\mu)$, $\mu \in (0, \infty)$. We know that $\hat{\mu}_{\text{MLE}}(x) = \bar{x}$ and $I_x = n/\bar{x}$. Let's look at the distribution of $Z = \sqrt{I_X}(\hat{\mu}_{\text{MLE}}(X) - \mu_0)$ for $\mu_0 = 1$, $n = 50$.

```
get.z <- function(x, mu.0, n){
  mu.mle <- mean(x)
  I.x <- n / mean(x)
  return(sqrt(I.x) * (mu.mle - mu.0))
}
M <- 10000
n <- 50
mu.0 <- 1
z.samp <- replicate(M, get.z(rpois(n, mu.0), mu.0, n))
hist(z.samp, freq = FALSE, col = "gray", border = "white")
z.grid <- seq(-5, 5, .1)
lines(z.grid, dnorm(z.grid))
```

> Task 1. Comment on how good the $\mathsf{Normal}(0,1)$ approximation is. Is there a particular feature of the histogram that the approximation fails to capture the most? Explain.
> Repeat the above experiment with $n = 100$ and $n = 200$. Comment how the approximation improves.

> Task 2. Now repeat the experiments with $\mu_0 = 5$ (for $n = 50, 100, 200$). Comment on any differences you notice from the case of $\mu_0 = 1$.

TASK 3. Repeat the above experiment for the model $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Exponential}(\lambda)$, $\lambda \in (0, \infty)$, for $n = 50, 100, 200$ and $\lambda_0 = 1, 5$. Comment on how well the $\mathsf{Normal}(0,1)$ distribution match the histogram of $Z = \sqrt{I_X}(\hat{\lambda}_{\text{MLE}}(X) - \lambda_0)$. Recall, $\hat{\lambda}_{\text{MLE}}(x) = 1/\bar{x}$, $I_x = n\bar{x}^2$.

Asymptotic Normality of Median

Many other estimators exhibit asymptotic normality. For example, if $X_1, \cdots, X_n \overset{\text{IID}}{\sim} g(x_i|\mu)$, where $g(x_i|\mu) = g_0(x_i - \mu)$ for a pdf g_0 symmetric around zero, then X_{med} is approximately $\mathsf{Normal}(\mu, \frac{1}{4ng_0(0)^2})$, i.e., $Z = \frac{X_{\text{med}} - \mu}{2g_0(0)\sqrt{n}}$ is approximately $\mathsf{Normal}(0, 1)$.

TASK 4. Demonstrate asymptotic normality of the median for

1. $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(0,1)$ $[g_0(y) = (2\pi)^{-1/2}\exp(-y^2/2)]$
2. $X_i \overset{\text{IID}}{\sim} \text{Laplace}(0,1)$ $[g_0(y) = (1/2)\exp(-|y|)]$
3. $X_i \overset{\text{IID}}{\sim} \text{Logis}(0,1)$ $[g_0(y) = \exp(-y)/\{1 + \exp(-y)\}^2]$

To demonstrate for the normal you can use the following code

```
M <- 10000
n <- 50
mu <- 0
g0.term <- 1 / sqrt(2 * pi)
sigma <- 1 / (2 * g0.term * sqrt(n))
z.samp <- replicate(M, (median(rnorm(n)) - mu) / sigma)
hist(z.samp, freq = FALSE, col = "gray", border = "white")
z.grid <- seq(-5, 5, .1)
lines(z.grid, dnorm(z.grid))
```

For the logistic case, use the above code but replace `rnorm` with `rlogis` [and use the appropriate formula for `g0.term`]. And for the Laplace model, use `rlap` instead, where the code is given below:

```
rlap <- function(n, mu = 0, sigma = 1){
  u <- runif(n)
  return(mu + sigma * sign(u - .5) * log(2 * pmin(u, 1 - u)))
}
```

Comparing various asymptotic CIs

For the model $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Poisson}(\mu)$, $\mu \in (0, \infty)$, the ML asymptotically 95%-CI for μ is

$$B_{1.96}(x) = \hat{\mu}_{\text{MLE}}(x) \mp 1.96/\sqrt{I_x} = \bar{x} \mp 1.96\sqrt{\bar{x}/n},$$

i.e., the confidence coefficient $\gamma(B_{1.96})$ is approximately 0.95 for large n (gets closer with larger values of n). Two other asymptotically 95%-CIs for μ are:

$$C_{1.96}(x) = \bar{x} \mp 1.96 \frac{s_x}{\sqrt{n}}$$

$$D_{1.96}(x) = \left(\bar{x} + \frac{1.96^2}{2n} - \frac{1.96}{\sqrt{n}} \left\{ \frac{1.96^2}{4n} + \bar{x} \right\}^{1/2}, \bar{x} + \frac{1.96^2}{2n} + \frac{1.96}{\sqrt{n}} \left\{ \frac{1.96^2}{4n} + \bar{x} \right\}^{1/2} \right)$$

Our goal is to compare these three interval procedures by their coverage and length. Here's how to get these characteristics for B_c for a given n at a fixed μ_0.

```
get.B <- function(x){
  x.bar <- mean(x)
  n <- length(x)
  B <- x.bar + c(-1, 1) * 1.96 * sqrt(x.bar / n)
  return(B)
}

features <- function(x, mu.0, int.proc){
  Int <- int.proc(x)
  Int.length <- Int[2] - Int[1]
  Int.inclusion <- (Int[1] <= mu.0) & (mu.0 <= Int[2])
  return(c(length = Int.length, coverage = Int.inclusion))
}

M <- 10000
n <- 10
mu.0 <- 1
options(digits = 2)
rowMeans(replicate(M, features(rpois(n, mu.0), mu.0, get.B)))
```

Task 5. Complete the comparison by first writing codes for `get.C` and `get.D`. Repeat this for $n = 10, 20, 50$ and $\mu_0 = 1, 5$.