STAT 510: Homework 07

David Dalpiaz

Due: Monday, April 4, 11:59 PM

Exercise 1 (Poisson Fisher Information)

Let $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. Find the method of moments estimator of λ. Find the maximum likelihood estimator of λ. Find the Fisher information $I(\lambda)$.

Solution

To obtain the **method of moments estimator**, we first find the first moment.

$$\mathbb{E}[X] = \lambda$$

Equating this with the first sample moment, we have our estimator.

$$\boxed{\hat{\lambda}_{\text{MOM}} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

To obtain the **maximum likelihood estimator**, we first obtain the log likelihood.

$$\log \mathcal{L}(\lambda) = \left(\sum_{i=1}^{n} x_i\right)\log(\lambda) - n\lambda - \sum_{i=1}^{n}\log(x_i!)$$

Maximizing this function gives us the MLE.

$$\boxed{\hat{\lambda}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

Lastly, we find the Fisher information.

$$I(\lambda) = -\mathbb{E}_{\lambda}\left[\frac{\partial^2 \log f(x;\lambda)}{\partial \lambda^2}\right] = \mathbb{E}\left[\frac{X}{\lambda^2}\right] = \boxed{\frac{1}{\lambda}}$$

Exercise 2 (Fisher Information Matrix)

Let $X_1, X_2, \ldots, X_n \sim \text{N}(\mu, \sigma^2)$. Find $I_n(\mu, \sigma)$.

Solution

First, recall the density of a normal distribution

$$f(x_i;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}.$$

Next, we find the log-likelihood.

$$\ell_n(\mu,\sigma) = \sum_{i=1}^{n} \log(f(X_i;\mu,\sigma)) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{2\sigma^2}$$

Now, we take the necessary derivatives.

$$\frac{\partial \ell_n(\mu,\sigma)}{\partial \mu} = \frac{\sum_{i=1}^{n}(X_i-\mu)}{\sigma^2}$$

$$\frac{\partial \ell_n(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^3}$$

$$\frac{\partial^2 \ell_n(\mu,\sigma)}{\partial^2 \mu} = -\frac{n}{\sigma^2}$$

$$\frac{\partial^2 \ell_n(\mu,\sigma)}{\partial \mu \partial \sigma} = -2\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sigma^3}$$

$$\frac{\partial^2 \ell_n(\mu,\sigma)}{\partial \sigma \partial \mu} = -2\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sigma^3}$$

$$\frac{\partial^2 \ell_n(\mu,\sigma)}{\partial^2 \sigma} = \frac{n}{\sigma^2} - 3\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^4}$$

Lastly, by arranging the relevant elements, we arrive at the request information matrix.

$$\boxed{I_n(\mu,\sigma) = -\mathbb{E}\begin{bmatrix} \frac{\partial^2 \ell_n(\mu,\sigma)}{\partial^2 \mu} & \frac{\partial^2 \ell_n(\mu,\sigma)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \ell_n(\mu,\sigma)}{\partial \sigma \partial \mu} & \frac{\partial^2 \ell_n(\mu,\sigma)}{\partial^2 \sigma} \end{bmatrix} = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}}$$

Exercise 3 (Exponential MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{Exponential}(\lambda)$. That is

$$f(x) = \lambda e^{-\lambda x}.$$

Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for λ.

Solution

First, we obtain the log-likelihood.

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

$$\frac{\partial}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

After completing the appropriate setups to maximize, we obtain the MLE.

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Next, we obtain the score function.

$$s(x; \lambda) = \frac{\partial}{\partial \lambda} \log\left(\lambda e^{-\lambda x}\right) = \frac{1}{\lambda} - x$$

Before calculating the Fisher information, we take another derivative.

$$\frac{\partial^2}{\partial \lambda^2} \log\left(\lambda e^{-\lambda x}\right) = -\frac{1}{\lambda^2}$$

Now we can find the Fisher information for a single observation.

$$I(\lambda) = -\mathbb{E}\left[-\frac{1}{\lambda^2}\right] = \frac{1}{\lambda^2}$$

Because our sample is independent and identically distributed, we can easily obtain the information for the entire sample.

$$I_n(\lambda) = \frac{n}{\lambda^2}$$

With this, we can obtain the estimated standard error of our estimator.

$$\hat{\text{se}}\left(\hat{\lambda}\right) = \sqrt{\frac{1}{I_n\left(\hat{\lambda}\right)}} = \frac{\sqrt{n}}{\sum_{i=1}^{n} x_i}$$

Finally, we can obtain an expression for an approximate 95% confidence interval.

$$\boxed{\frac{n}{\sum_{i=1}^{n} x_i} \pm 2 \cdot \frac{\sqrt{n}}{\sum_{i=1}^{n} x_i}}$$

Exercise 4 (Exponential MLE, Continued)

Define $\phi = \log(\lambda)$. Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for ϕ.

```
set.seed(42)
exp_data = rexp(n = 100, rate = 0.5)
```

Using the data stored in `exp_data`, calculate an approximate 95% confidence interval for λ two ways:

- Using the interval from Exercise 3.
- Using the interval from this exercise, transformed back to the λ scale.

Solution

To find the MLE of ϕ, we simply appeal to the equivariance property of the MLE.

$$\hat{\phi} = \log\left(\hat{\lambda}\right) = \log\left(\frac{n}{\sum_{i=1}^{n} x_i}\right)$$

To obtain the standard error, we use the delta method.

$$\hat{\text{se}}\left(\hat{\phi}\right) = \left|g'(\hat{\lambda})\right| \cdot \hat{\text{se}}\left(\hat{\lambda}\right) = \frac{\sum_{i=1}^{n} x_i}{n} \cdot \frac{\sqrt{n}}{\sum_{i=1}^{n} x_i} = \frac{1}{\sqrt{n}}$$

Thus, we can obtain an expression for an approximate 95% confidence interval for ϕ.

$$\log\left(\frac{n}{\sum_{i=1}^{n} x_i}\right) \pm 2 \cdot \frac{1}{\sqrt{n}}$$

```
calc_ci_lambda = function(data) {
  n = length(data)
  n / sum(data) + c(-2, 2) * sqrt(n) / sum(data)
}
```

```
calc_ci_phi = function(data) {
  n = length(data)
  log(n / sum(data)) + c(-2, 2) / sqrt(n)
}
```

Using the two functions above that calculate approximate 95% confidence intervals for λ and ϕ respectively, we calculate the two requested intervals.

```
# interval for lambda
calc_ci_lambda(exp_data)
```

```
## [1] 0.3557737 0.5336605
```

```
# interval for phi, converted to lambda scale
exp(calc_ci_phi(exp_data))
```

```
## [1] 0.3641036 0.5431787
```

Exercise 5 (Another MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{N}(\theta, 1)$. Define

$$Y_i = \begin{cases} 1 & \text{if } X_i > 0 \\ 0 & \text{if } X_i \leq 0. \end{cases}$$

Let $\phi = \mathbb{P}(Y_1 = 1)$.

Use the MLE, $\hat{\phi}$, and its standard error to derive an expression for an approximate 95% confidence interval for ϕ.

Solution

First, we note that the MLE for θ is well known.

$$\hat{\theta} = \bar{X}$$

Note that

$$\hat{\mathrm{se}}\left(\hat{\theta}\right) = \frac{1}{\sqrt{n}}$$

Next, we obtain an expression for ϕ that involves the CDF of a standard normal, $\Phi(z)$. Note that the use of ϕ as a parameter is a coincidence here.

$$\phi = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_i > 0) = \mathbb{P}\left(\frac{X_i - \theta}{1} > \frac{0 - \theta}{1}\right) = \mathbb{P}(Z > -\theta) = \mathbb{P}(Z < \theta) = \Phi(\theta)$$

Using this expression together with the equivariance property of the MLE, we obtain the MLE for ϕ.

$$\hat{\phi} = \Phi(\hat{\theta}) = \Phi(\bar{X})$$

Now using the delta method, we obtain the estimated standard error of this MLE.

$$\hat{\mathrm{se}}\left(\hat{\phi}\right) = \left|\Phi'(\hat{\theta})\right| \cdot \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{2\pi}\sqrt{n}} e^{-\frac{1}{2}\bar{x}^2}$$

Lastly, we tie these results together to obtain an approximate 95% confidence interval.

$$\boxed{\Phi(\bar{x}) \pm 2 \cdot \frac{1}{\sqrt{2\pi}\sqrt{n}} e^{-\frac{1}{2}\bar{x}^2}}$$

Exercise 6 (Asymptotic Relative Efficiency)

Continuing the setup from Exercise 5, now define

$$\tilde{\phi} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

Find the asymptotic relative efficiency of $\tilde{\phi}$ to $\hat{\phi}$. Your answer will be a function of θ. Provide the value of of θ and the associated asymptotic relative efficiency for the value of θ that gives the largest asymptotic relative efficiency.

Solution

First, note that $Y_i \sim \text{Bernoulli}(\phi)$.

$$\mathbb{V}[Y_i] = \phi \cdot (1 - \phi) = \Phi(\theta) \cdot (1 - \Phi(\theta))$$

$$\mathbb{V}\left[\tilde{\phi}\right] = \frac{\phi \cdot (1 - \phi)}{n} = \frac{\Phi(\theta) \cdot (1 - \Phi(\theta))}{n}$$

From the previous exercise, we can also obtain the variance of $\hat{\phi}$.

$$\mathbb{V}\left[\hat{\phi}\right] = \frac{\left(\Phi'(\theta)\right)^2}{n}$$

Recall that

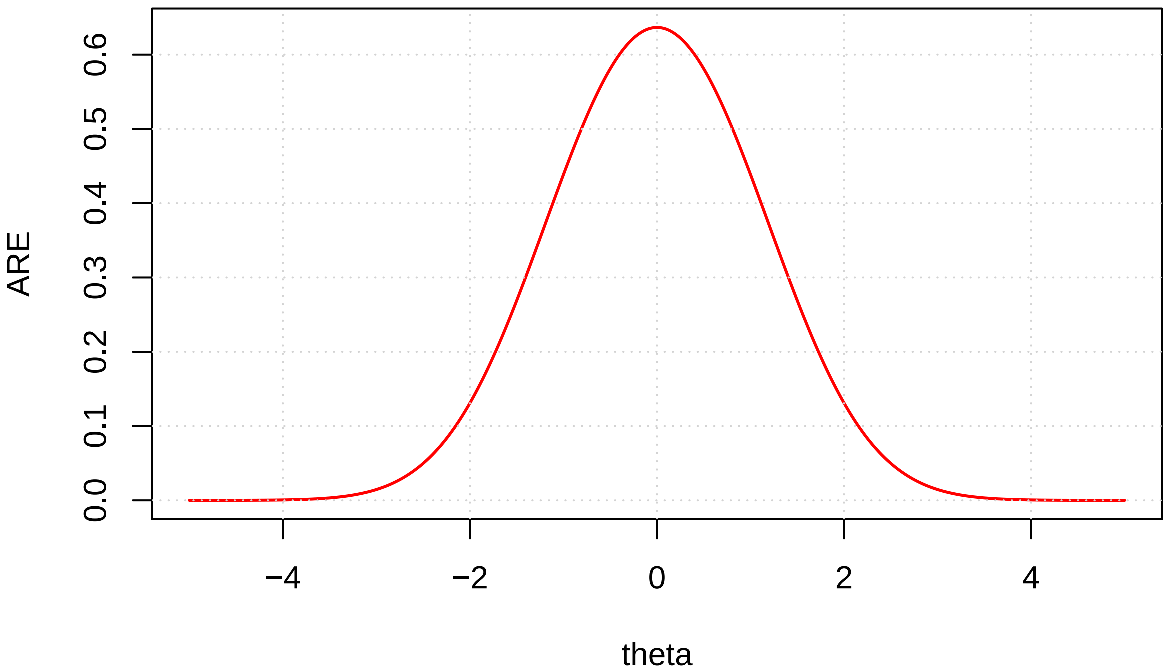
$$\Phi'(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{x}^2}$$

Now, we finally investigate the ARE.

$$\boxed{\text{ARE}\left(\tilde{\phi}, \hat{\phi}\right) = \frac{\mathbb{V}\left[\hat{\phi}\right]}{\mathbb{V}\left[\tilde{\phi}\right]} = \frac{\left(\Phi'(\theta)\right)^2}{\Phi(\theta) \cdot (1 - \Phi(\theta))}}$$

```r
calc_are = function(theta) {
  top = dnorm(theta) ^ 2
  bottom = pnorm(0, mean = theta) * pnorm(0, mean = theta, lower.tail = FALSE)
  return(top / bottom)
}
```

```r
x = seq(-5, 5, by = 0.01)
plot(x, calc_are(x), type = "l", col = "red", lwd = 1.5,
     xlab = "theta", ylab = "ARE")
grid()
```



By observing the plot above, it is clear that the ARE is maximized at a $\boxed{\theta = 0}$. This corresponds to $\boxed{\text{ARE} = 2/\pi}$.

```r
calc_are(0)
```

```
## [1] 0.6366198
```

```r
2 / pi
```

```
## [1] 0.6366198
```

Exercise 7 (Comparing Two Groups)

Suppose n_1 people are given treatment 1 and n_2 people are given treatment 2. Let X_1 be the number of people on treatment 1 who respond favorably to the treatment and let X_2 be the number of people on treatment 2 who respond favorably.

Assume $X_1 \sim \text{Binomial}(n_1, p_1)$ and $X_2 \sim \text{Binomial}(n_2, p_2)$.

Let $\phi = p_1 - p_2$.

Use the MLE, $\hat{\phi}$, and its standard error to derive an expression for an approximate 90% confidence interval for ϕ. To arrive at the standard error, first find $I(p_1, p_2)$ and then apply the delta method.

Solution

First, we obtain the likelihood. Note that c_i is a placeholder for a term that does not depend on the parameters.

$$\mathcal{L}(p_1, p_2) = \left[c_1 \cdot p_1^{x_1}(1 - p_1)^{n_1 - x_1}\right] \cdot \left[c_2 \cdot p_2^{x_2}(1 - p_2)^{n_2 - x_2}\right]$$

We then find the log-likelihood, as well as derivatives with respect to each of the parameters.

$$\log \mathcal{L}(p_1, p_2) = \log c_1 + x_1 \log p_1 + (n_1 - x_1)\log(1 - p_1) + \log c_2 + x_2 \log p_2 + (n_2 - x_2)\log(1 - p_2)$$

$$\frac{\partial}{\partial p_1} \log \mathcal{L}(p_1, p_2) = \frac{x_1}{p_1} - \frac{n_1 - x_1}{1 - p_1}$$

$$\frac{\partial}{\partial p_2} \log \mathcal{L}(p_1, p_2) = \frac{x_2}{p_2} - \frac{n_2 - x_2}{1 - p_2}$$

We also note the second order derivatives as they will be needed later.

$$\frac{\partial^2}{\partial p_1^2} \log \mathcal{L}(p_1, p_2) = -\frac{x_1}{p_1^2} - \frac{n_1 - x_1}{(1 - p_1)^2}$$

$$\frac{\partial^2}{\partial p_2^2} \log \mathcal{L}(p_1, p_2) = -\frac{x_2}{p_2^2} - \frac{n_2 - x_2}{(1 - p_2)^2}$$

$$\frac{\partial^2}{\partial p_1 \partial p_2} \log \mathcal{L}(p_1, p_2) = \frac{\partial^2}{\partial p_2 \partial p_1} \log \mathcal{L}(p_1, p_2) = 0$$

Returning to the first order derivatives, after setting them equal to zero and solving, we obtain the MLE for p_1 and p_2.

$$\hat{p}_1 = \frac{x_1}{n_1}$$

$$\hat{p}_2 = \frac{x_2}{n_2}$$

So the MLE of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2$.

Next, we obtain the individual elements of the information matrix.

$$H_{11} = \frac{\partial^2}{\partial p_1^2} \log \mathcal{L}(p_1, p_2) = -\frac{x_1}{p_1^2} - \frac{n_1 - x_1}{(1 - p_1)^2}$$

$$H_{22} = \frac{\partial^2}{\partial p_2^2} \log \mathcal{L}(p_1, p_2) = -\frac{x_2}{p_2^2} - \frac{n_2 - x_2}{(1 - p_2)^2}$$

$$H_{12} = \frac{\partial^2}{\partial p_1 \partial p_2} \log \mathcal{L}(p_1, p_2) = \frac{\partial^2}{\partial p_2 \partial p_1} \log \mathcal{L}(p_1, p_2) = H_{21} = 0$$

$$\mathbb{E}[H_{11}] = -\frac{n_1}{p_1(1 - p_1)}$$

$$\mathbb{E}[H_{22}] = -\frac{n_2}{p_2(1 - p_2)}$$

$$\mathbb{E}[H_{12}] = \mathbb{E}[H_{21}] = 0$$

So, finally, we have the full information matrix.

$$I_n(p_1, p_2) = \begin{bmatrix} \frac{n_1}{p_1(1-p_1)} & 0 \\ 0 & \frac{n_2}{p_2(1-p_2)} \end{bmatrix}$$

Now we work to apply the delta method.

$$J_n(p_1, p_2) = I_n^{-1}(p_1, p_2) = \begin{bmatrix} \frac{p_1(1-p_1)}{n_1} & 0 \\ 0 & \frac{p_2(1-p_2)}{n_2} \end{bmatrix}$$

$$\phi = g(p_1, p_2) = p_1 - p_2$$

$$\nabla g = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\hat{\text{se}}(\hat{\phi}) = \sqrt{(\hat{\nabla} g)^T \hat{J}_n (\hat{\nabla} g)} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Thus, finally we arrive at an approx 90% confidence interval.

$$\boxed{\hat{p}_1 - \hat{p}_2 \ \pm \ 1.645 \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

Exercise 8 (Comparing Standard Errors)

Continue with the setup from Exercise 7. Given:

- $n_1 = n_2 = 200$
- $X_1 = 160$
- $X_2 = 148$

Compare 90% confidence interval for ϕ using standard errors from Exercise 7 and the parametric bootstrap.

Solution

```r
# data
n_1 = 200
n_2 = 200
x_1 = 160
x_2 = 148
```

```r
# estimates
p_1_hat = x_1 / n_1
p_2_hat = x_2 / n_2
est = p_1_hat - p_2_hat
se_hat = sqrt(p_1_hat * (1 - p_1_hat) / n_1 + p_2_hat * (1 - p_2_hat) / n_2)
```

```r
# critical value
crit = qnorm(0.95)
```

```r
# interval via large sample theory and delta method
est + c(-1, 1) * crit * se_hat
```

```
## [1] -0.009044678  0.129044678
```

```r
# function for parametric bootstrap
boot_est = function() {
  x_1 = rbinom(n = 1, size = 200, prob = 160 / 200)
  x_2 = rbinom(n = 1, size = 200, prob = 148 / 200)

  x_1 / 200 - x_2 / 200
}
```

```r
# estimate the standard error via bootstrap
se_boot = sd(replicate(n = 5000, boot_est()))
```

```r
# interval via parametric bootstrap
est + c(-1, 1) * crit * se_boot
```

```
## [1] -0.00876081  0.12876081
```

Note that the interval obtained via the parametric bootstrap is ever so slightly larger.

Exercise 9 (Geometric MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{Geometric}(\pi)$.

Use the MLE, $\hat{\pi}$, and its standard error to derive an expression for an approximate 95% confidence interval for π.

Solution

First, we note the density of a geometric random variable.

$$f(x; \pi) = (1 - \pi)^{x-1} \cdot \pi$$

Next, we obtain the likelihood and the log-likelihood.

$$\mathcal{L}(\pi) = \prod_{i=1}^{n} (1 - \pi)^{x_i - 1} \cdot \pi$$

$$\log \mathcal{L}(\pi) = \left(\sum_{i=1}^{n} x_i - n \right) \log(1 - \pi) + n \log \pi$$

We then proceed to find the maximum.

$$\frac{\partial}{\partial \pi} \log \mathcal{L}(\pi) = -\frac{\sum_{i=1}^{n} x_i - n}{1 - \pi} + \frac{n}{\pi}$$

After equating the above with zero and solving, we obtain the MLE for π.

$$\hat{\pi} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Now we work towards finding the standard error.

$$\log f(x; \pi) = (x - 1) \log(1 - \pi) + \log \pi$$

$$\frac{\partial}{\partial \pi} \log f(x;\pi) = \frac{1-x}{1-\pi} + \frac{1}{\pi}$$

$$\frac{\partial^2}{\partial \pi^2} \log f(x;\pi) = \frac{1-x}{(1-\pi)^2} - \frac{1}{\pi^2}$$

Using the above, we first find the information of a single observation.

$$I(\pi) = -\mathbb{E}\left[\frac{\partial^2}{\partial \pi^2} \log f(x;\pi)\right] = \mathbb{E}\left[\frac{x-1}{(1-\pi)^2} + \frac{1}{\pi^2}\right] = \frac{1}{\pi^2(1-\pi)}$$

Since we are dealing with independent and identically distributed random variables, it is easy to obtain the information provided by the entire sample.

$$I_n(\pi) = n \cdot I(\pi) = \frac{n}{\pi^2(1-\pi)}$$

Finally, we obtain the estimated standard error.

$$\hat{\text{se}}(\hat{\pi}) = \sqrt{\frac{\hat{\pi}^2(1-\hat{\pi})}{n}}$$

And thus we obtain an approximate 95% confidence interval for π.

$$\hat{\pi} \ \pm \ 2\sqrt{\frac{\hat{\pi}^2(1-\hat{\pi})}{n}}$$

Recall that $\hat{\pi}$ is the MLE of *pi*. The expression for this MLE is repeated below.

$$\hat{\pi} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Exercise 10 (Geometric MLE, Continued)

Define $\psi = \text{logit}(\pi)$. Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for ψ.

```
set.seed(42)
geom_data = rgeom(n = 100, prob = 0.2)
```

Using the data stored in `geom_data`, calculate an approximate 95% confidence interval for π two ways:

- Using the interval from Exercise 9.
- Using the interval from this exercise, transformed back to the π scale.

Solution

$$\psi = \text{logit}(\pi) = \log \frac{\pi}{1-\pi}$$

Last exercise, we found the MLE of π and its estimated standard error.

$$\hat{\pi} = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\hat{\text{se}}(\hat{\pi}) = \sqrt{\frac{\hat{\pi}^2(1-\hat{\pi})}{n}}$$

The MLE of ψ is found via equivariance.

$$\hat{\psi} = g(\hat{\pi}) = \text{logit}(\hat{\pi}) = \log \frac{\hat{\pi}}{1-\hat{\pi}}$$

We now use the delta method to find the standard error.

$$g'(\hat{\pi}) = \frac{1}{\hat{\pi}(1-\hat{\pi})}$$

$$\hat{\text{se}}\left(\hat{\psi}\right) = \left|\frac{1}{\hat{\pi}(1-\hat{\pi})}\right| \sqrt{\frac{\hat{\pi}^2(1-\hat{\pi})}{n}} = \frac{1}{\sqrt{n(1-\hat{\pi})}}$$

Now we obtain the requested interval.

$$\log \frac{\hat{\pi}}{1-\hat{\pi}} \ \pm \ 2 \cdot \frac{1}{\sqrt{n(1-\hat{\pi})}}$$

```
calc_ci_pi = function(data) {
  n = length(data)
  pi_hat = n / sum(data)
  pi_hat + c(-2, 2) * sqrt( pi_hat ^ 2 * (1 - pi_hat) / n)
}
```

```
calc_ci_psi = function(data) {
  n = length(data)
  pi_hat = n / sum(data)
  boot::logit(pi_hat) + c(-2, 2) / sqrt(n * (1 - pi_hat))
}
```

```
# shift data to match parameterization use in previous problem
geom_data = geom_data + 1
```

```
# interval for pi
calc_ci_pi(geom_data)
```

```
## [1] 0.1746277 0.2500007
```

```
# interval for psi, converted to pi scale
boot::inv.logit(calc_ci_psi(geom_data))
```

```
## [1] 0.1770622 0.2524319
```

Exercise 11 (Rao-Blackwellization)

Let $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for λ. Consider two estimators:

1. $\hat{\lambda}_1 = X_1$
2. $\hat{\lambda}_2$ which is the results of applying Rao-Blackwell to $\hat{\lambda}_1 = X_1$ with $\sum_{i=1}^n X_i$.

Show that $\hat{\lambda}_2$ has a smaller MSE than $\hat{\lambda}_1$.

Solution

$$f(X_1, \ldots, X_n; \lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i} e^{-\lambda}}{X_i!} = \left(e^{\sum_{i=1}^{n} X_i \log \lambda - n\lambda}\right) \cdot \left(\prod_{i=1}^{n} X_i!\right)$$

Thus using the Factorization Theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for λ.

Then using Rao-Blackwell, we obtain $\hat{\lambda}_2$.

$$\hat{\lambda}_2 = \mathbb{E}\left[X_1 \mid \sum_{i=1}^{n} X_i\right]$$

Note that both estimators are unbiased.

$$\mathbb{E}\left[\hat{\lambda}_1\right] = \lambda$$

To obtain this result for $\hat{\lambda}_2$ we apply the law of iterated expectation.

$$\mathbb{E}\left[\hat{\lambda}_2\right] = \mathbb{E}\left[\mathbb{E}\left[X_1 \mid \sum_{i=1}^{n} X_i\right]\right] = \mathbb{E}\left[X_1\right] = \lambda$$

Thus, for both, the MSE of the estimator is simply its variance. Recalling the law of iterated variances, note that for random variables X and Y we have

$$\mathbb{V}[Y] = \mathbb{E}[\mathbb{V}[Y \mid X]] + \mathbb{V}[\mathbb{E}[Y \mid X]]$$

Thus, we have

$$\mathbb{V}\left[\hat{\lambda}_2\right] = \mathbb{V}\left[\mathbb{E}\left[X_1 \mid \sum_{i=1}^{n} X_i\right]\right] = \mathbb{V}[X_1] - \mathbb{E}\left[\mathbb{V}\left[X_1 \mid \sum_{i=1}^{n} X_i\right]\right] \leq \mathbb{V}[X_1] = \mathbb{V}\left[\hat{\lambda}_1\right]$$

The inequality above holds by definition of variance.

$$\mathbb{V}\left[X_1 \mid \sum_{i=1}^{n} X_i\right] \geq 0.$$

Thus we have shown

$$\mathrm{MSE}\left[\hat{\lambda}_2\right] \leq \mathrm{MSE}\left[\hat{\lambda}_1\right].$$