2. Delta method and asymptotic variances

(a) (Optional)

0 points possible (ungraded)

In this problem, you are going to compute the **asymptotic variance** of some estimators. Recall that the asymptotic variance of an estimator $\hat{\theta}$ for a parameter θ is defined as $V(\hat{\theta})$, if

$$\sqrt{n}\,(\hat{\theta}-\theta) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, V(\hat{\theta})\right).$$

The arguments that we use to establish asymptotic normality are often the same in our setups, namely the Law of Large Numbers, the Central Limit Theorem, and the Delta Method. First, we review the assumptions and statements of those theorems:

Let $X_1, X_2, \ldots,$ be random variables. The (weak) Law of Large Numbers says that under suitable assumptions, with

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

we have

$$\overline{X}_n \xrightarrow{\mathbf{P}} \mathbb{E}[X_1].$$

What are the assumptions we need for the weak Law of Large Numbers? (Choose all that apply.)

- [x] $\mathbb{E}[|X_i|] < \infty$ for all i ✔
- [ ] $\mathsf{Var}(X_i) < \infty$ for all i
- [x] $X_1, X_2, \ldots$ independent ✔
- [ ] There exists $M > 0$ such that $|X_i| \le M$ for all i
- [ ] $|X_i| \ge |X_{i+1}|$ almost surely for all i
- [x] $X_1, X_2, \ldots$ identically distributed ✔

✔

The Central Limit Theorem states that under some assumptions, there is a V such that

$$\sqrt{n}\,(\overline{X}_n - \mathbb{E}[X_1]) \xrightarrow{(\mathrm{D})} \mathcal{N}(0, V).$$

What are the assumptions we need for the Central Limit Theorem? Pick all that apply.

- [x] $\mathbb{E}[|X_i|] < \infty$ for all i ✔

- [x] $\text{Var}(X_i) < \infty$ for all i ✔
- [x] $X_1, X_2, \ldots$ independent ✔
- [ ] There exists $M > 0$ such that $|X_i| \leq M$ for all i
- [ ] $|X_i| \geq |X_{i+1}|$ almost surely for all i
- [x] $X_1, X_2, \ldots$ identically distributed ✔

✔

The Delta Method gives us a way to control the asymptotic variance of a transformation of a random variable. Let $\theta \in \mathbb{R}$ be a parameter and $Z_n \in \mathbb{R}$ be a sequence of random variables that satisfies

$$\sqrt{n}\left(Z_n - \theta\right) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, V\right)$$

for some $V > 0$.

Given a function $g : \Omega \subseteq \mathbb{R} \to \mathbb{R}$,

$$\sqrt{n}\left(g\left(Z_n\right) - g\left(\theta\right)\right) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, W\right).$$

for some $W > 0$.

Pick the following assumptions and conditions that apply to the Delta Method as stated in class:

- [ ] g is monotonically increasing
- [x] g is continuously differentiable at θ ✔
- [x] $W = g'(\theta)^2 V$ ✔
- [ ] $W = g(\theta)^2 V$
- [ ] $W = |g'(\theta)| V$

✔

Solution:

For the weak Law of Large Numbers to apply, we need that the X_i are independent and identically distributed (although there exist weaker versions of it). Moreover, the limit expectation needs to actually exist, i.e. $\mathbb{E}\left[|X_i|\right] < \infty$.

For the Central Limit Theorem, we have the same requirements, and on top of that, we need the variance to be finite, i.e. $\text{Var}(X_i) < \infty$.

For the Delta Method, we need that g is continuously differentiable at θ and the correct asymptotic variance is given by $W = g'(\theta)^2 V$.

Note: There is also a multivariate version of the Delta Method, which we will discuss later in this course.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

Instructions:
Now, in each of the following questions, argue that both proposed estimators are consistent and asymptotically normal. Then, give their asymptotic variances and decide if one of them is always bigger than the other.

(b)

1/2 points (graded)
Argue that the proposed estimators $\widehat{\lambda}$ and $\tilde{\lambda}$ below are both consistent and asymptotically normal. Then, give their asymptotic variances $V(\widehat{\lambda})$ and $V(\tilde{\lambda})$, and decide if one of them is always bigger than the other.

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \textsf{Poiss}(\lambda)$, for some $\lambda > 0$. Let $\hat{\lambda} = \overline{X}_n$ and $\tilde{\lambda} = -\ln(\overline{Y}_n)$, where $Y_i = \mathbf{1}\{X_i = 0\}, i = 1, \ldots, n$.

$V(\widehat{\lambda}) =$ lambda ✔ **Answer:** lambda

λ

$V(\tilde{\lambda}) =$ exp(-1*lambda)*(1-exp(✖ **Answer:** exp(lambda) - 1

$\frac{\exp(-1\cdot\lambda)\cdot(1-\exp(-1\cdot\lambda))}{\lambda^2}$

STANDARD NOTATION

Solution:

For $\widehat{\lambda}$, By the Law of Large Numbers,

$$\overline{X}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[X_1] = \lambda.$$

By the Central Limit Theorem,

$$\sqrt{n}(\overline{X}_n - \lambda) \sim \mathcal{N}(0, \textsf{Var}(X_1)) = \mathcal{N}(0, \lambda),$$

hence

$$V(\widehat{\lambda}) = \lambda.$$

For $\tilde{\lambda}$, first observe that by the Law of Large Numbers,

$$\overline{Y}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[Y_1] = \mathbf{P}(X_1 = 0) = \exp(-\lambda),$$

so with $g(t) = -\log(t)$

$$\tilde{\lambda} = g(\overline{Y}_n) \xrightarrow[n\to\infty]{\mathbf{P}} g(\exp(-\lambda)) = \lambda.$$

The Central Limit Theorem yields

$$\sqrt{n}(\overline{Y}_n - \mathbb{E}[Y_1]) \xrightarrow[n\to\infty]{(D)} \mathcal{N}(0, \textsf{Var}(Y_1)) = \mathcal{N}(0, \exp(-\lambda)(1 - \exp(-\lambda)),$$

where we used the formula $\mathsf{Var}(Z) = p(1-p)$ if $Z \sim \mathsf{Be}(p)$. In order to apply the Delta Method for the above $g(t)$, we compute

$$g'(t) = -\frac{1}{t}, \quad g'(\exp(-\lambda)) = -\exp(\lambda),$$

which results in

$$\sqrt{n}(\tilde{\lambda} - \lambda) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}(0, \exp(\lambda) - 1).$$

Moreover, by the series expansion for the exponential,

$$\exp(\lambda) - 1 = \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} > \lambda, \quad \text{for all } \lambda > 0,$$

so $V(\widehat{\lambda}) < V(\tilde{\lambda})$ for all λ.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(c)

1/3 points (graded)

As above, argue that both proposed estimators $\widehat{\lambda}$ and $\tilde{\lambda}$ are consistent and asymptotically normal. Then, give their asymptotic variances $V(\widehat{\lambda})$ and $V(\tilde{\lambda})$, and decide if one of them is always bigger than the other.

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathsf{Exp}(\lambda)$, for some $\lambda > 0$. Let $\widehat{\lambda} = \frac{1}{\overline{X}_n}$ and $\tilde{\lambda} = -\ln(\overline{Y}_n)$, where $Y_i = \mathbf{1}\{X_i > 1\}, i = 1, \ldots, n$.

$V(\widehat{\lambda}) =$ lambda^2 ✔ **Answer:** lambda^2

λ^2

$V(\tilde{\lambda}) =$ exp(-1*lambda)*(1-exp(✖ **Answer:** exp(lambda) - 1

$\frac{\exp(-1\cdot\lambda)\cdot(1-\exp(-1\cdot\lambda))}{\lambda^2}$

- $V(\widehat{\lambda}) > V(\tilde{\lambda})$ for all λ.
- $V(\widehat{\lambda}) < V(\tilde{\lambda})$ for all λ. ✔
- $V(\widehat{\lambda}) = V(\tilde{\lambda})$ for all λ.
- There exists λ_1 such that $V(\widehat{\lambda}) > V(\tilde{\lambda})$ and λ_2 such that $V(\widehat{\lambda}) < V(\tilde{\lambda})$ ✖

STANDARD NOTATION

Solution:

For $\widehat{\lambda}$, by the Law of Large Numbers,

$$\overline{X}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[X_1] = \frac{1}{\lambda}.$$

With $g(t) = 1/t$, we have that

$$\widehat{\lambda} \xrightarrow[n\to\infty]{\mathbf{P}} \frac{1}{\mathbb{E}[X_1]} = \lambda.$$

By the Central Limit Theorem,

$$\sqrt{n}\,(\overline{X}_n - \frac{1}{\lambda}) \sim \mathcal{N}(0, \mathsf{Var}(X_1)) = \mathcal{N}\left(0, \frac{1}{\lambda^2}\right).$$

The fact that

$$g'(t) = -\frac{1}{t^2}$$

together with the Delta Method then yields

$$V(\widehat{\lambda}) = \lambda^2.$$

For $\tilde{\lambda}$, first observe that it is the average of Bernoulli variables, and by the Law of Large Numbers,

$$\overline{Y}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[Y_1] = \mathbf{P}(X_1 > 1) = \exp(-\lambda),$$

so with $\tilde{g}(t) = -\log(t)$

$$\tilde{\lambda} = \tilde{g}(\overline{Y}_n) \xrightarrow[n\to\infty]{\mathbf{P}} g(\exp(-\lambda)) = \lambda.$$

The Central Limit Theorem yields

$$\sqrt{n}\,(\overline{Y}_n - \mathbb{E}[Y_1]) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}(0, \mathsf{Var}(Y_1)) = \mathcal{N}(0, \exp(-\lambda)(1 - \exp(-\lambda)).$$

In order to apply the Delta Method for the above $\tilde{g}(t)$, we compute

$$\tilde{g}'(t) = -\frac{1}{t}, \quad \tilde{g}'(\exp(-\lambda)) = -\exp(\lambda),$$

which results in

$$\sqrt{n}\,(\tilde{\lambda} - \lambda) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}(0, \exp(\lambda) - 1).$$

In order to compare these two asymptotic variances, first observe that similar to part (b),

$$\exp(\lambda) - 1 \geq \lambda, \quad \text{for all } \lambda > 0,$$

and since $\lambda^2 < \lambda$ for $\lambda \in (0,1)$, we have

$$\exp(\lambda) - 1 \geq \lambda^2, \quad \text{for } \lambda \in (0,1).$$

Moreover,

$$\exp(1) - 1 = e > 1 = 1^2,$$

and

$$\frac{d}{d\lambda}(\exp(\lambda) - 1) = \exp(\lambda), \quad \frac{d}{d\lambda}\lambda^2 = 2\lambda,$$

so that

$$\frac{d}{d\lambda}(\exp(\lambda) - 1) = \exp(\lambda) \geq 1 + \lambda + \frac{\lambda^2}{2} > 2\lambda = \frac{d}{d\lambda}\lambda^2, \quad \text{for all } \lambda > 0,$$

which can be checked by the quadratic formula. This means that for $\lambda \geq 1$,

$$\exp(\lambda) - 1 = e + \int_1^\lambda \exp(t)\, dt > 1 + \int_1^\lambda 2t\, dt = \lambda^2.$$

Hence, the asymptotic variance of $\widehat{\lambda}$ is always lower than that of $\tilde{\lambda}$.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(d) (Optional)

0 points possible (ungraded)

Ungrading Note: *This question needs techniques that you will only learn in a later unit, and we will revisit this problem in a later homework. For now, this is ungraded. We are sorry for the oversight on our part, but hope that the time you have spent on this undoable question (for now) will still be worthwhile in the long run.*

As above, argue that both proposed estimators $\widehat{\sigma^2}$ and $\widetilde{\sigma^2}$ are consistent and asymptotically normal. Then, give their asymptotic variances $V\left(\widehat{\sigma^2}\right)$ and $V\left(\widetilde{\sigma^2}\right)$ and decide if one of them is always bigger than the other.

Let $X_1, \ldots, X_n \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$, for some $\sigma^2 > 0$. Let

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^n X_i^2, \quad \text{and} \quad \widetilde{\sigma^2} = \frac{1}{n}\sum_{i=1}^n \left(X_i - \overline{X}_n\right)^2.$$

$V\left(\widehat{\sigma^2}\right) =$ sigma^4 ✖

σ^4

$V(\widetilde{\sigma^2}) =$ 2*sigma^2 ✖

$2 \cdot \sigma^2$

STANDARD NOTATION

提交 你已经尝试了2次（总共可以尝试2次）

(e)

3/3 points (graded)

As above, argue that both proposed estimators $\hat{p}$, $\tilde{p}$, and are consistent and asymptotically normal. Then, give their asymptotic variances $V(\hat{p})$ and $V(\tilde{p})$ and decide if one of them is always bigger than the other.

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim}$ Geom (p), for some $p \in (0,1)$. That means that

$$\mathbf{P}(X_1 = k) = p(1-p)^{k-1}, \quad \text{for } k = 1, 2, \ldots.$$

Let

$$\hat{p} \;=\; \frac{1}{\overline{X}_n},$$

and $\tilde{p}$ be the **number of ones in the sample divided by** n.

$V(\hat{p}) =$ p^2*(1-p) ✔ **Answer:** p^2*(1-p)

$p^2 \cdot (1-p)$

$V(\tilde{p}) =$ p*(1-p) ✔ **Answer:** p*(1-p)

$p \cdot (1-p)$

- ○ $V(\hat{p}) > V(\tilde{p})$ for all p.
- ◉ $V(\hat{p}) < V(\tilde{p})$ for all p. ✔
- ○ $V(\hat{p}) = V(\tilde{p})$ for all p.
- ○ There exists p_1 such that $V(\hat{p}) > V(\tilde{p})$ and p_2 such that $V(\hat{p}) < V(\tilde{p})$

STANDARD NOTATION

Solution:

By the Law of Large Numbers,

$$\overline{X}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[X_1] = \frac{1}{p}.$$

Setting

$$g(t) = \frac{1}{t},$$

we obtain consistency of $\hat{p} = g\left(\overline{X}_n\right)$, i.e.,

$$\hat{p} = g\left(\overline{X}_n\right) \xrightarrow[n\to\infty]{\mathbf{P}} g\left(\mathbb{E}\left[X_1\right]\right) = p.$$

By the Central Limit Theorem,

$$\sqrt{n}\left(\overline{X}_n - \frac{1}{p}\right) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, \mathsf{Var}\left(X_1\right)\right) = \mathcal{N}\left(0, \frac{1-p}{p^2}\right),$$

and hence by the Delta Method, together with

$$g'\left(\frac{1}{p}\right)^2 = \left(-p^2\right)^2 = p^4,$$

we end up with

$$\sqrt{n}\left(\hat{p} - p\right) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, p^2\left(1-p\right)\right),$$

so

$$V\left(\hat{p}\right) = p^2\left(1-p\right).$$

For $\tilde{p}$, note that we can write it as

$$\tilde{p} = \overline{Y}_n, \quad \text{where } Y_i = \mathbf{1}\{X_i = 1\},$$

so it is again an average over Bernoulli variables. The Law of Large Numbers gives

$$\overline{Y}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}\left[Y_1\right] = \mathbf{P}\left(X_1 = 1\right) = p,$$

while the Central Limit Theorem yields

$$\sqrt{n}\left(\overline{Y}_n - p\right) \xrightarrow[n\to\infty]{(\mathrm{D})} \mathcal{N}\left(0, \mathsf{Var}\left(Y_1\right)\right) = \mathcal{N}\left(0, p\left(1-p\right)\right).$$

Since $p^2 < p$ for $p \in (0,1)$,

$$V\left(\hat{p}\right) < V\left(\tilde{p}\right).$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

讨论

显示讨论