Homework 5

(a)

1 point possible (graded)

$X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = X^2$. Please enter in terms of μ and σ .

$\text{Cov}(X, Y) =$ **Answer:** 2*mu*sigma^2

STANDARD NOTATION

Solution:

The definition for the covariance of two random variables: $\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$. An alternative form for the covariance is $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. This form is easier to work with to calculate covariances compared to the original definition.

$\mathbb{E}[X^2] = \sigma^2 + \mu^2$, $\mathbb{E}[X^3] = \mu^3 + 3\mu\sigma^2$.

$$\begin{aligned} \text{Cov}(X, X^2) &= \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2] \\ &= \mu^3 + 3\mu\sigma^2 - \mu(\mu^2 + \sigma^2) \\ &= 2\mu\sigma^2 \end{aligned}$$

(b)

1 point possible (graded)

X , Y have the joint probability density function $f(x, y) = 1$, $0 < x < 1$, $x < y < x + 1$. Please enter a number.

$\text{Cov}(X, Y) =$ **Answer:** 1/12

Solution:

$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, so we need to find out the expectations of X , Y , and XY . From the joint distribution, we can derive the marginal distribution: $f_X(x) = \int_x^{x+1} 1\, dy = y|_x^{x+1} = 1$, $x \in (0, 1)$ and the conditional distribution $f(y|x) = \frac{f(x,y)}{f(x)} = 1$, $y \in (x, x + 1)$.

On one hand, we have $\mathbb{E}[X] = \frac{1}{2}$: since $f_X(x)$ doesn't depend on x, this describes the density of a uniform random variable over $[0, 1]$. On the other hand, for the mean of Y:

$$\begin{aligned} \mathbb{E}[Y|X] &= \int_x^{x+1} y\, dy \\ &= \left.\frac{y^2}{2}\right|_x^{x+1} \\ &= \frac{2x + 1}{2} \end{aligned}$$

According to the law of iterated expectations,

$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y|X]] \\ &= \mathbb{E}[\frac{2X+1}{2}] \\ &= \int_0^1 \frac{2x+1}{2}\,dx \\ &= 1\end{aligned}$$

$$\begin{aligned}\mathbb{E}[XY] &= \int_0^1 x\left[\int_x^{x+1} y\,dy\right]dx \\ &= \int_0^1 x\frac{y^2}{2}\Big|_x^{x+1}\,dx \\ &= \int_0^1 \frac{2x^2+x}{2}\,dx \\ &= \frac{7}{12}\end{aligned}$$

$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{7}{12} - \frac{1}{2}\times 1 = \frac{1}{12}$

(c)

1 point possible (graded)

$X \sim f(x) = \frac{1}{2b}e^{-|x|/b},\ x \in \mathbb{R},\ b > 0$ and $Y = \text{sign}(X)$

$\text{Cov}(X,Y) =$ **Answer:** b

Solution:

By symmetry, $\mathbb{E}[X] = \int_{-\infty}^{\infty} \frac{x}{2b}e^{-|x|/b}\,dx = 0$. $\mathbb{E}[Y] = (-1)\cdot P(X<0) + 1\cdot P(X>0) = -\frac{1}{2} + \frac{1}{2} = 0$

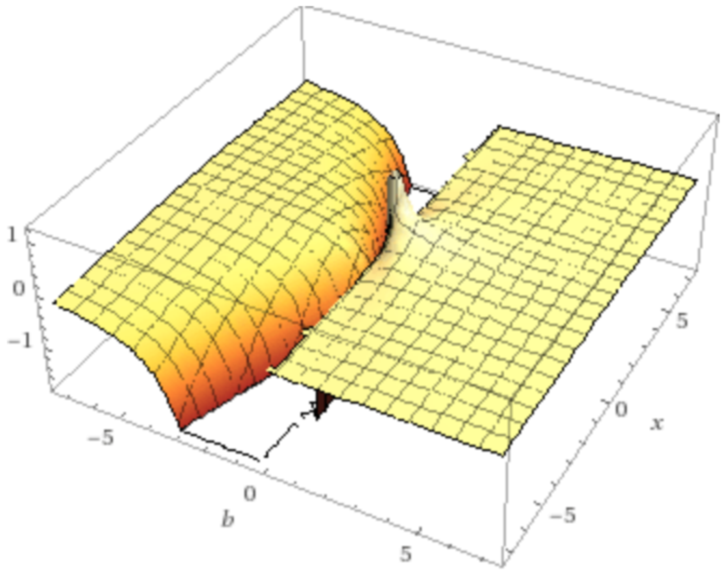
$$\begin{aligned}\text{Cov}(X,Y) = \mathbb{E}[XY] &= \int_{-\infty}^{\infty} \frac{x\cdot\text{sign}(x)}{2b}e^{-|x|/b}\,dx \\ &= \int_0^{\infty} \frac{x}{b}e^{-x/b}\,dx\end{aligned}$$

We can think of this as the expectation of an exponential random variable Z with parameter $\frac{1}{b}$.
$\int_0^{\infty} \frac{x}{b}e^{-x/b}\,dx = \mathbb{E}[Z] = b$, where $Z \sim \text{Exp}(\frac{1}{b})$.

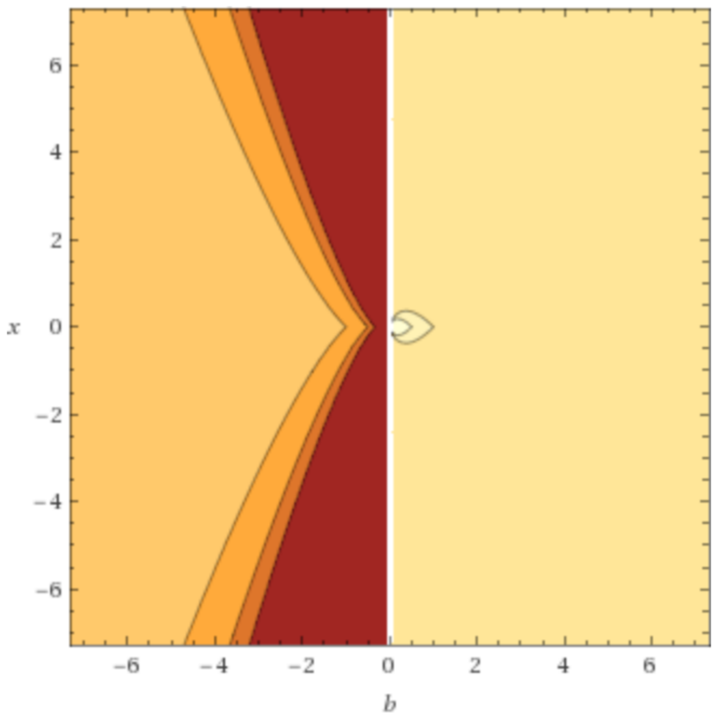
symmetric around x=0?

plot	$\frac{1}{2b}\exp\left(-\frac{\|x\|}{b}\right)$

3D plot:



Contour plot:



(d)

1 point possible (graded)

$X \sim$ Unif $(0, 1)$ and given $X = x,\ Y \sim$ Unif $(x, 1)$

Cov $(X, Y) =$

Answer: 1/24

Solution:

$\mathbb{E}[X] = \frac{1}{2}$
$\mathbb{E}[Y|X] = \frac{X+1}{2}$
According to the law of iterated expectations, $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[\frac{X+1}{2}] = \int_0^1 \frac{x+1}{2}\,dx = \frac{3}{4}$

$$f(x, y) = f(y|x)\, f(x) = \frac{1}{1-x}$$

$$\begin{aligned}
\mathbb{E}[XY] &= \int_0^1 \int_x^1 \frac{1}{1-x} \cdot xy\, dydx \\
&= \int_0^1 \frac{x}{1-x} \cdot \frac{y^2}{2}\Big|_x^1 dx \\
&= \int_0^1 \frac{x}{1-x}(\frac{1}{2} - \frac{x^2}{2})\ dx \\
&= \frac{1}{2}\int_0^1 (x + x^2)\ dx \\
&= \frac{5}{12}
\end{aligned}$$

Cov $(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = \frac{5}{12} - \frac{1}{2} \times \frac{3}{4} = \frac{1}{24}$

(f)

1 point possible (graded)

$X + Y$ and $X - Y$, where X and Y are independent $\mathcal{N}(\mu, \sigma^2)$.

Cov $(X + Y, X - Y) =$

Answer: 0

Solution:

$$\begin{aligned}
\text{Cov}(X+Y, X-Y) &= \mathbb{E}[(X+Y)(X-Y)] - \mathbb{E}[X+Y]\,\mathbb{E}[X-Y] \\
&= \mathbb{E}[X^2] - \mathbb{E}[Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])(\mathbb{E}[X] - \mathbb{E}[Y]) \\
&= (\sigma^2 + \mu^2) - (\sigma^2 + \mu^2) - ((\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2) \\
&= 0
\end{aligned}$$

2. A Simple Singular Covariance Matrix

Homework due Jul 1, 2020 08:59 JST *Past Due*

Bookmark this page

Suppose $\mathbf{X}$ is a random vector, where $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})^T$, with mean $\mathbf{0}$ and covariance matrix $\mathbf{v}\mathbf{v}^T$, for some vector $\mathbf{v} \in \mathbb{R}^d$.

(a)

1 point possible (graded)

If $d > 1$, is the covariance matrix $\mathbf{v}\mathbf{v}^T$ invertible?

Hint: Compute the determinant for the case $d = 2$. That result will generalize to higher dimension.

- ○ $\mathbf{v}\mathbf{v}^T$ is invertible.
- ● $\mathbf{v}\mathbf{v}^T$ is **not** invertible. ✔

Solution:

For $d > 1$, the matrix $\mathbf{v}\mathbf{v}^T$ where $\mathbf{v}$ is a vector in $\mathbb{R}^d$ is not invertible. To get an intuition, we start with an example in 2 dimensions:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \implies \mathbf{v}\mathbf{v}^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

The matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is not invertible. One way to see this is that its determinant is $1\,(0) - (0)\,(0) = 0$. Another way to see it is that for any 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In fact, the above argument works in general after a change of variables. Given $\mathbf{v} \in \mathbb{R}^d$, change coordinates of $\mathbb{R}^d$ so that the first axis points in the direction of $\mathbf{v}$ (and so that $\mathbf{v}$ has unit length). In this new coordinate system, $\mathbf{v}$ can be rewritten as $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, and the matrix

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & 0 \end{pmatrix}$$

is not invertible because no $d \times d$ matrix when multiplied by it will give the identity matrix.

v is a vector (here matrix of size $d \times 1$). Hence, vv^T is a matrix of size $d \times d$.

In the first question, we are asked to either prove a statement (vv^T is invertible) or to find a counterexample for this statement. The answer discusses the counterexample when $d = 2$ and $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Note that $v^T v$ is quite a different thing. It is a matrix of size 1×1 (a scalar) and it is equal to the square of the norm $\|v\|$ (L_2 norm) of the vector (see here for vector norms). For instance, for the case $v = \begin{pmatrix} a \\ b \end{pmatrix}$, we have

$$v^T v = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a^2 + b^2 = \|v\|^2$$

Also, if you look at the general case, see figure below from this question, you can understand why it is by construction a covariance matrix.

so $a^T b$ is equivalent to $a \cdot b$, while

$$aa^T = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} a_1^2 & a_1 a_2 & \cdots & a_1 a_n \\ a_2 a_1 & a_2^2 & \cdots & a_2 a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n a_1 & a_n a_2 & \cdots & a_n^2 \end{bmatrix}.$$

The matrix presented in the solution for problem (a) is just an example for the vector: $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Let's pick another vector (for simplicity in 2D) : $\mathbf{v_1} = \begin{pmatrix} a \\ b \end{pmatrix}$.

Then, $\mathbf{v_1}^T = \begin{pmatrix} a & b \end{pmatrix}$

$$\mathbf{v_1 v_1}^T = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}$$

Then, matrix is invertible IFF the determinant $\neq 0$

$\det(\mathbf{v_1 v_1}^T) = a^2 \cdot b^2 - ab \cdot ab = 0$. Hence, $\mathbf{v_1 v_1}^T$ is **not** invertible.

(b)

1 point possible (graded)

Let $\mathbf{u}$ be a vector in $\mathbb{R}^d$ such that $\mathbf{u} \perp \mathbf{v}$, i.e. $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u} = 0$.

Find the variance of $\mathbf{u}^T\mathbf{X}$.

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of a vector $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\mathsf{Var}\,(\mathbf{u}^T\mathbf{X}) =$ **Answer:** 0

STANDARD NOTATION

Solution:

Given two vectors $\mathbf{u}, \mathbf{X} \in \mathbb{R}^d$, the inner product $\mathbf{u}^T\mathbf{X}$ is a scalar, and its variance is also a scalar. Using the covariance matrix formula, we get

$$\begin{aligned}\mathsf{Var}\,(\mathbf{u}^T\mathbf{X}) = \mathsf{Cov}\,(\mathbf{u}^T\mathbf{X}) &= \mathbf{u}^T\mathsf{Cov}\,(\mathbf{X})\,\mathbf{u}\\ &= \mathbf{u}^T\,(\mathbf{v}\mathbf{v}^T)\,\mathbf{u}\\ &= (\mathbf{u}^T\mathbf{v})\,(\mathbf{v}^T\mathbf{u}) = 0.\end{aligned}$$

(c)

1 point possible (graded)

Let $\overline{\mathbf{v}} = \dfrac{\mathbf{v}}{||\mathbf{v}||}$ (i.e., $\overline{\mathbf{v}}$ is the normalized version of $\mathbf{v}$). What is the variance of $\overline{\mathbf{v}}^T\mathbf{X}$?

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\mathsf{Var}\,(\overline{\mathbf{v}}^T\mathbf{X}) =$ **Answer:** norm(v)^2

STANDARD NOTATION

Solution:

Similarly

$$\begin{aligned}\mathsf{Var}\,(\overline{\mathbf{v}}^T\mathbf{X}) = \mathsf{Cov}\,(\overline{\mathbf{v}}^T\mathbf{X}) &= \overline{\mathbf{v}}^T\mathsf{Cov}\,(X)\,\overline{\mathbf{v}}\\ &= \left(\frac{\mathbf{v}}{||\mathbf{v}||}\right)^T(\mathbf{v}\mathbf{v}^T)\left(\frac{\mathbf{v}}{||\mathbf{v}||}\right)\\ &= \frac{(\mathbf{v}^T\mathbf{v})\,(\mathbf{v}^T\mathbf{v})}{||\mathbf{v}||^2} = ||\mathbf{v}||^2.\end{aligned}$$

(d)

1 point possible (graded)

Suppose we observe n independent copies of $\mathbf{X}$ and call them $\mathbf{X}_1, \ldots, \mathbf{X}_n$. What is the asymptotic distribution of $\overline{\mathbf{X}}_n = \frac{\sum_{i=1}^n \mathbf{X}_i}{n}$? (Select all that apply.)

- [ ] $\sqrt{n}\,(\overline{\mathbf{X}}_n - \mathbf{0}) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(\mathbf{0}, \mathbf{I}_d)$ where $\mathbf{I}_d$ is the identity matrix in R^d
- [ ] $\sqrt{n}\,(\overline{\mathbf{X}}_n - \mathbf{0}) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(\mathbf{0}, \mathbf{v}\mathbf{v}^T)$ ✔
- [ ] $\sqrt{n}\,(\overline{\mathbf{X}}_n - \mathbf{0}) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(\mathbf{0}, \|\mathbf{v}\|^2)$

Note on notation: In the choices above, $\mathcal{N}$ denotes a multivariate Gaussian distribution. In lecture and elsewhere, a multivariate Gaussian distribution in d dimension is also sometimes denoted with an extra subscript by $\mathcal{N}_d$. To be accurate, read the dimension from the arguments, i.e. the mean and the covariance matrix.

Solution:

By multivariate CLT,

$$\sqrt{n}\,(\overline{\mathbf{X}}_n - \mathbf{0}) \xrightarrow[n\to\infty]{(d)} \mathcal{N}\,(\mathbf{0}, \mathbf{v}\mathbf{v}^T)$$

However, $\mathbf{v}\mathbf{v}^T$ is not invertible, so the pdf of $\mathcal{N}\,(\mathbf{0}, \mathbf{v}\mathbf{v}^T)$ is not given by the usual formula that involves the inverse of the determinant of the covariance matrix of the multivariate Gaussian variable.

(e)

2 points possible (graded)

Let $\mathbf{Y}_i = \overline{\mathbf{v}}\,(\overline{\mathbf{v}}^T \mathbf{X}_i)$, or equivalently $\overline{\mathbf{v}}\,(\overline{\mathbf{v}} \cdot \mathbf{X}_i) = (\overline{\mathbf{v}} \cdot \mathbf{X}_i)\,\overline{\mathbf{v}}$, where $\overline{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ is the same as in part (c).

We will compare the asymptotic distribution of $\overline{\mathbf{X}}_n$ you obtain in part (d) to the asymptotic distribution of $\overline{\mathbf{Y}}_n$ where $\overline{\mathbf{Y}}_n = \frac{\sum_i^n \mathbf{Y}_i}{n}$.

What is the expectation $\mathbb{E}\,[\mathbf{Y}_i]$ of $\mathbf{Y}_i$?
(Choose all that apply.)

- [x] $\overline{\mathbf{v}}\overline{\mathbf{v}}^T\,\mathbb{E}\,[\mathbf{X}_i]$ ✔
- [x] $\mathbf{0}$ (the zero vector in $\mathbb{R}^d$) ✔
- [ ] 0 (the real number zero)
- [ ] $\overline{\mathbf{v}}^T\,\mathbf{v}$

Find the covariance matrix $\Sigma_{\mathbf{Y}_i}$ of $\mathbf{Y}_i$ in terms of the vector $\mathbf{v}$.

(If applicable, enter **trans(v)** for the transpose $\mathbf{v}^T$ of $\mathbf{v}$, and **norm(v)** for the norm $||\mathbf{v}||$ of a vector $\mathbf{v}$.)

$\Sigma_{\mathbf{Y}_i} =$ **Answer:** v*trans(v)

(There is no answer box for the following question.)

Notice that $\mathbf{Y}_i$ is a scalar multiple of the vector $\mathbf{v}$ and hence lies on the same line as $\mathbf{v}$ no matter what value $\mathbf{X}_i$ takes. (Specifically, $\mathbf{Y}_i = (\overline{\mathbf{v}}^T \mathbf{X}_i)\,\overline{\mathbf{v}}$ is the projection of $\mathbf{X}_i$ onto the vector $\mathbf{v}$.) Use your answers for $\mathbb{E}[\mathbf{Y}_i]$ and $\Sigma_{\mathbf{Y}_i}$ to find the asymptotic distribution of $\overline{\mathbf{Y}}_n$. Compare this with the asymptotic distribution of $\overline{\mathbf{X}}_n$ from the previous part. What can you conclude about the asymptotic distribution of $\overline{\mathbf{X}}_n$?

3. Asymptotic Variance

Homework due Jul 1, 2020 08:59 JST *Past Due*

Bookmark this page

a)

2 points possible (graded)

Note: *This question is the ungraded problem from homework 2.*

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

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

Argue that both proposed estimators $\widehat{\sigma^2}$ and $\widetilde{\sigma^2}$ below are consistent and asymptotically normal.

Then, give their asymptotic variances $V(\widehat{\sigma^2})$ and $V(\widetilde{\sigma^2})$ and decide if one of them is always bigger than the other.

Hint: Use the multivariate Delta method. Also see Recitation 5 *Inference for the Variance of a Gaussian distribution*.

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

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

Solution:

Note that

$$\widehat{\sigma^2} = \overline{Y}_n, \qquad \text{for } Y_i = X_i^2.$$

By the Law of Large Numbers,

$$\overline{Y}_n \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[Y_1] = \sigma^2.$$

By the Central Limit Theorem,

$$\sqrt{n}(\overline{Y}_n - \sigma^2) \sim \mathcal{N}(0, \mathsf{Var}(Y_1)) = \mathcal{N}(0, 2(\sigma^2)^2),$$

hence

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

For $\widetilde{\sigma^2}$, first observe that we can write it as

$$\begin{aligned}\widetilde{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2\overline{X}_n X_i + \overline{X}_n^2) \\ &= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2\right) - \overline{X}_n^2 = \widehat{\sigma^2} - \overline{X}_n^2.\end{aligned}$$

Again, by the Law of Large Numbers,

$$\overline{X}_n^2 \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[X_1]^2 = 0,$$

so

$$\widetilde{\sigma^2} = \widehat{\sigma^2} - \overline{X}_n^2 \xrightarrow[n\to\infty]{\mathbf{P}} \sigma^2.$$

Now, we can consider $\widetilde{\sigma^2}$ as

$$\widetilde{\sigma^2} = g\left(\frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix}X_i \\ X_i^2\end{pmatrix}\right),$$

where

$$g(x, y) = y - x^2.$$

By the above, we have a multidimensional Central Limit Theorem for the first and second moments of a Gaussian together,

$$\sqrt{n}\left[\begin{pmatrix} \overline{X}_n \\ \overline{Y}_n \end{pmatrix} - \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix}\right] \xrightarrow[n\to\infty]{(D)} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2(\sigma^2)^2 \end{pmatrix}\right),$$

where the 0 s on the diagonal come from the fact that

$$\mathbb{E}[X_i \times X_i^2] = \mathbb{E}[X_i^3] = 0.$$

Now, apply the multidimensional Delta Method, computing

$$Dg(x, y) = \begin{pmatrix} -2x & 1 \end{pmatrix},$$

to obtain

$$\begin{aligned} \sqrt{n}(\widetilde{\sigma^2} - \sigma^2) &\xrightarrow[n\to\infty]{(D)} \mathcal{N}\left(0, Dg(0, \sigma^2)\begin{pmatrix} \sigma^2 & 0 \\ 0 & 2(\sigma^2)^2 \end{pmatrix} Dg(0, \sigma^2)^\top\right) \\ &= \mathcal{N}\left(0, \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} \sigma^2 & 0 \\ 0 & 2(\sigma^2)^2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \mathcal{N}(0, 2(\sigma^2)^2). \end{aligned}$$

Combined, we see that both estimators have the same asymptotic variance.

4. Maximum Likelihood Estimator for Curved Gaussian

Homework due Jul 1, 2020 08:59 JST *Past Due*

Bookmark this page

(a)

1 point possible (graded)

Note: *To avoid too much double jeopardy, the solution to part (a) will be available once you have either answered it correctly or reached the maximum number of attempts.*

Let $X_1, \ldots, X_n$ be n i.i.d. random variables with distribution $\mathcal{N}(\theta, \theta)$ for some unknown $\theta > 0$.

Compute the maximum likelihood estimator $\hat{\theta}$ for θ in terms of the sample averages of the linear and quadratic means, i.e. $\overline{X}_n$ and $\overline{X_n^2}$.

(Enter **barX_n** for $\overline{X}_n$ and **bar(X_n^2)** for $\overline{X_n^2}$. Note that **barX_n^2** represents $(\overline{X}_n)^2$, and is **not** equal to **bar(X_n^2)** with the brackets.

$\hat{\theta} =$ **Answer:** (sqrt(4 * bar(X_n^2) + 1) - 1)/2

Solution:

To compute the maximum likelihood estimator, we write the log likelihood and maximize it by setting its derivative to zero. First,

$$\begin{aligned}
\ell_n(\theta) &= \sum_{i=1}^{n} \log\left[\frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{(X_i-\theta)^2}{2\theta}\right)\right] \\
&= -\frac{n}{2}(\log(2) + \log(\pi) + \log(\theta)) - \sum_{i=1}^{n} \frac{(X_i-\theta)^2}{2\theta} \\
&= -\frac{n}{2}(\log(2) + \log(\pi) + \log(\theta)) - \sum_{i=1}^{n} \left[\frac{1}{2\theta}X_i^2 - X_i + \frac{1}{2}\theta\right].
\end{aligned}$$

Differentiating yields

$$\frac{d}{d\theta}\ell(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} X_i^2 - \frac{n}{2},$$

which we set to zero to obtain the equation

$$\hat{\theta}^2 + \hat{\theta} - \frac{1}{n}\sum_{i=1}^{n} X_i^2 = 0.$$

Employing the quadratic formula and picking the result that gives a positive $\hat{\theta}$ then leads to

$$\hat{\theta} = -\frac{1}{2} + \frac{1}{2}\sqrt{4\overline{X_n^2} + 1}.$$

(b)

4 points possible (graded)

We want to compute the asymptotic variance of $\hat{\theta}$ via two methods.

In this problem, we apply the Central Limit Theorem and the 1-dimensional Delta Method. We will compare this with the approach using the Fisher information next week.

First, compute the limit and asymptotic variance of $\overline{X_n^2}$.

The limit to which $\overline{X_n^2}$ converges in probability, also known as its **$\mathbf{P}$-limit**, is

$\overline{X_n^2} \xrightarrow[n\to\infty]{\mathbf{P}}$

Answer: theta + theta^2

The asymptotic variance $V(\overline{X_n^2})$ of $\overline{X_n^2}$, which is equal to $\mathsf{Var}(X_1^2)$, is

$V(\overline{X_n^2}) = \mathsf{Var}(X_1^2) =$

Answer: 2*theta^2*(2*theta + 1)

Now, write $\hat{\theta}$ as the function of $\overline{X_n^2}$ you found in part (a),

$$\hat{\theta} = g(\overline{X_n^2})$$

and give its first derivative, $g'(x)$,

$g'(x) =$

Answer: 1/sqrt(4*x+1)

What can you conclude about the asymptotic variance $V(\hat{\theta})$ of $\hat{\theta}$?

$V(\hat{\theta}) =$

Answer: 2*theta^2/(2*theta + 1)

STANDARD NOTATION

Solution:

First, by the Law of Large Numbers,

$$\overline{X_n^2} \xrightarrow[n\to\infty]{\mathbf{P}} \mathbb{E}[X_1^2] = \mathsf{Var}(X_1) + \mathbb{E}[X_1]^2 = \theta + \theta^2.$$

Its asymptotic variance can be found by the Central Limit Theorem that gives us

$$\sqrt{n}\left(\overline{X_n^2} - (\theta+\theta^2)\right) \xrightarrow[n\to\infty]{(D)} \mathcal{N}\left(0, \mathsf{Var}(X_1^2)\right),$$

and

$$\begin{aligned} \mathsf{Var}(X_1^2) = &\ \mathbb{E}[X_1^4] - \mathbb{E}[X_1^2]^2 \\ = &\ \mathbb{E}[(\theta + \sqrt{\theta}Z)^4] - (\theta+\theta^2)^2 \\ = &\ \theta^4 + 4\theta^3\sqrt{\theta}\underbrace{\mathbb{E}[Z]}_{=0} + 6\theta^2\theta\underbrace{\mathbb{E}[Z^2]}_{=0} + 4\theta\sqrt{\theta}^3\underbrace{\mathbb{E}[Z^3]}_{=0} + \theta^2\underbrace{\mathbb{E}[Z^4]}_{=3} - \theta^4 - 2\theta^3 + \theta^2 \\ = &\ 2\theta^2(2\theta+1), \end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$ is a standard Normal variable.

From the previous part, we get

$$g(x) = \frac{1}{2}\left(\sqrt{4x+1} - 1\right),$$

so

$$g'(x) = \frac{1}{\sqrt{4x+1}}.$$

Finally, by the Delta Method,

$$\sqrt{n}\left(g\left(\overline{X_n^2}\right) - g(\theta+\theta^2)\right) \xrightarrow[n\to\infty]{(D)} \mathcal{N}\left(0, 2\theta^2(2\theta+1)\, g'(\theta+\theta^2)^2\right) = \mathcal{N}\left(0, \frac{2\theta^2}{2\theta+1}\right).$$