MATH 408 Homework 3

Qilin Ye

March 23, 2022

Problem 1

Suppose you know the following list of numbers is a random sample of size 20 from a Guassian with mean 1 and unknown variance $\sigma^2 > 0$.

$$2.0753 \quad 4.6678 \quad -3.5177 \quad 2.7243 \quad 1.6375 \quad -1.6154 \quad 0.1328 \quad 1.6852 \quad 8.1568 \quad 6.5389$$

$$-1.6998 \quad 7.0698 \quad 2.4508 \quad 0.8739 \quad 2.4295 \quad 0.5901 \quad 0.7517 \quad 3.9794 \quad 3.8181 \quad 3.8344.$$

(1) Using a method of moments estimator, estimate the value of σ^2 for this data.

(2) Denote your method of moments estimator for σ^2 as Z. Is Z unbiased?

(3) We know for sure that $\sigma^2 > 0$. Is it possible that Z could take negative values? If so, then perhaps Z is not the best way to estimate σ^2.

(4) The Delta method suggests that $1/Z$ could be a good estimate for $1/\sigma^2$. What estimate of $1/\sigma^2$ do you get from the data above? Is $\mathbb{E}|1/Z|$ finite? If not, then we cannot even compute the bias of this estimator.

(5) The Delta method also suggests that Z^2 could be a good estimate for σ^4. What estimate of σ^4 do you get from the data above? Is Z^2 an unbiased estimate of σ^4?

(6) Is Z^2 an asymptotically unbiased estimate of σ^4? That is, as the number of samples n goes to infinity, does $\mathbb{E}Z^2$ converge to σ^4?

Solution. (1) Since $\sigma^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}X^2 - 1$, using the method of moments,

$$\hat{\sigma}^2 = \frac{1}{20}\sum_{i=1}^{20} X_i^2 - 1 \approx 12.7452.$$

(2) Z is unbiased:

$$\mathbb{E}Z = \mathbb{E}(\frac{1}{20}\sum_{i=1}^{20} X_i^2 - 1) = \frac{1}{20}\mathbb{E}(\sum_{i=1}^{20} X_i^2) - 1 = \mathbb{E}X^2 - 1 = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \sigma^2.$$

(3) Yes, if each $X_i \in [0,1)$ then the corresponding $Z < 0$. This could well happen.

(4) From the data above we get $1/\hat{\sigma}^2 = 0.0785$. The expected value $\mathbb{E}|1/Z|$ is not finite. To see this, notice that $f_Z(z)$ is continuous on, say, $[1,1]$. By a well-known result in analysis,

$$\inf_{z\in[-1,1]} = \min_{z\in[-1,1]} = f_Z(\tilde{z}) \geqslant \epsilon > 0 \qquad \text{for some } z \in [-1,1] \text{ and } \epsilon > 0.$$

Since Z has codomain $[-1.\infty)$, we have

$$\mathbb{E}[1/Z] = \int_{-1}^{\infty} f_Z(z)\frac{1}{|z|}\,\mathrm{d}z \geqslant \int_{-1}^{1} f_Z(z)\frac{1}{|z|}\,\mathrm{d}z \geqslant 2\int_{0}^{1} \epsilon\frac{1}{z}\,\mathrm{d}z = \infty.$$

(5) The estimator we have is $\hat{\sigma}^4 = 162.4397$. It is biased:

$$\begin{aligned}
\mathbb{E}(Z^2) &= \mathbb{E}(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 1)^2 \\
&= \frac{1}{n^2}\mathbb{E}(\sum_{i=1}^{n} X_i^2)^2 - 2\cdot\frac{1}{n}\cdot\mathbb{E}(\sum_{i=1}^{n} X_i^2) + 1 \\
&= \frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^{n} X_i^4 + \sum_{i\neq j} X_i^2 X_j^2\right] - 2\mathbb{E}X^2 + 1 \\
&= \frac{1}{n^2}(n\mathbb{E}X^4 + (n^2-n)(\mathbb{E}X^2)^2) - 2\mathbb{E}X^2 + 1 \\
&= \frac{\mathbb{E}X^4}{n} + \frac{n-1\mathbb{E}(X^2)^2}{n} - 2\mathbb{E}(X^2) + 1,
\end{aligned}$$

whereas

$$\sigma^4 = (\mathbb{E}X^2 - 1)^2 = (\mathbb{E}X^2)^2 - 2\mathbb{E}X^2 + 1.$$

(6) Following the calculation in the previous part, letting $n \to \infty$, we obtain

$$\lim_{n\to\infty} E_n(Z^2) = \lim_{n\to\infty}(\ldots) = \sigma^4,$$

so Z is asymptotically unbiased.

Problem 2

Let $X, Y, Z : \Omega \to \mathbb{R}$ be discrete or continuous random variables. Let A be the range of Y. Define $g : A \to \mathbb{R}$ by $g(y) := \mathbb{E}(X \mid Y = y)$ for any $y \in A$. We then define the **conditional expectation** of X given Y, denoted $\mathbb{E}(X \mid Y)$, to be the random variable $g(Y)$.

(i) Let X, Y be random variables such that (X, Y) is uniform distributed on the triangle given by $\{(x, y) \in \mathbb{R}^2 : x \geqslant 0, y \geqslant 0, x + y \leqslant 1\}$. Show that $\mathbb{E}(X \mid Y) = (1 - Y)/2$.

(ii) Prove the following version of the Total Expectation Theorem:

$$\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(X).$$

(iii) Show the following

$$\mathbb{E}(X \mid X) = X \qquad \text{and} \qquad \mathbb{E}(X + Y \mid Z) = \mathbb{E}(X \mid Z) + \mathbb{E}(Y \mid Z).$$

(iv) If Z is independent of X and Y, show that

$$\mathbb{E}(X \mid Y, Z) = \mathbb{E}(X \mid Y).$$

Proof. (i) For $y \in [0,1]$, note that $X \mid Y = y$ is uniformly distributed on $[0, 1-y]$, so $\mathbb{E}(X \mid Y = y) = (1-y)/2$. Therefore by definition $\mathbb{E}(X \mid Y) = (1-Y)/2$.

(ii) For the continuous case:

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X \mid Y)) &= \int_{-\infty}^{\infty} f_Y(y)\mathbb{E}(X \mid Y = y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} f_Y(y)\left[\int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\,\mathrm{d}x\right]\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X|Y}(x,y) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \mathbb{E}X.
\end{aligned}$$

For the discrete case:

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(\sum_x x\mathbb{P}(X = x \mid Y = y)) &= \sum_y (\sum_x x\mathbb{P}(X = x \mid Y = y))\mathbb{P}(Y = y) \\
&= \sum_x \sum_y x\mathbb{P}(X = x \mid Y = y)\mathbb{P}(Y = y) \\
&= \sum_x x(\sum_y \mathbb{P}(X = x, Y = y)) = \sum_x x\mathbb{P}(X = x) = \mathbb{E}X.
\end{aligned}$$

(iii) The first claim is trivial, as $\mathbb{E}(X \mid X = x) = x$. The continuous case for the second equation:

$$\begin{aligned}
\mathbb{E}(X + Y \mid Z = z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f_{X+Y|Z}(x+y \mid z)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_{\mathbb{R}^2} x f_{X+Y|Z}(x+y \mid z)\,\mathrm{d}x\mathrm{d}y + \iint_{\mathbb{R}^2} y f_{X+Y|Z}(x+y \mid z)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X+Y|Z}(x+y \mid z)\mathrm{d}y\,\mathrm{d}x + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X+Y|Z}(x+y \mid z)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x f_{X|Z}(x \mid z)\,\mathrm{d}x + \int_{-\infty}^{\infty} y f_{Y|Z}(y \mid z)\,\mathrm{d}y = \mathbb{E}(X \mid Z = z) + \mathbb{E}(Y \mid Z = z).
\end{aligned}$$

The discrete case for the second equation:

$$\begin{aligned}
\mathbb{E}(X + Y \mid Z = z) &= \sum_x \sum_y (x+y)\mathbb{P}(X = x, Y = y \mid Z = z) \\
&= \sum_x x \sum_y \mathbb{P}(X = x, Y = y \mid Z = z) + \sum_y y \sum_x \mathbb{P}(X = x, Y = y \mid Z = z) \\
&= \sum_x \mathbb{P}(X = x \mid Z = z) + \sum_y y\mathbb{P}(Y = y \mid Z = z) = \mathbb{E}(X \mid Z = z) + \mathbb{E}(Y \mid Z = z).
\end{aligned}$$

(iv) If Z is independent of X and Y then (assuming they are continuous)

$$f_{X|(Y,Z)}(x \mid (y,z)) = \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)} = \frac{f_{X,Y}(x,y)}{f_Y(y)},$$

so

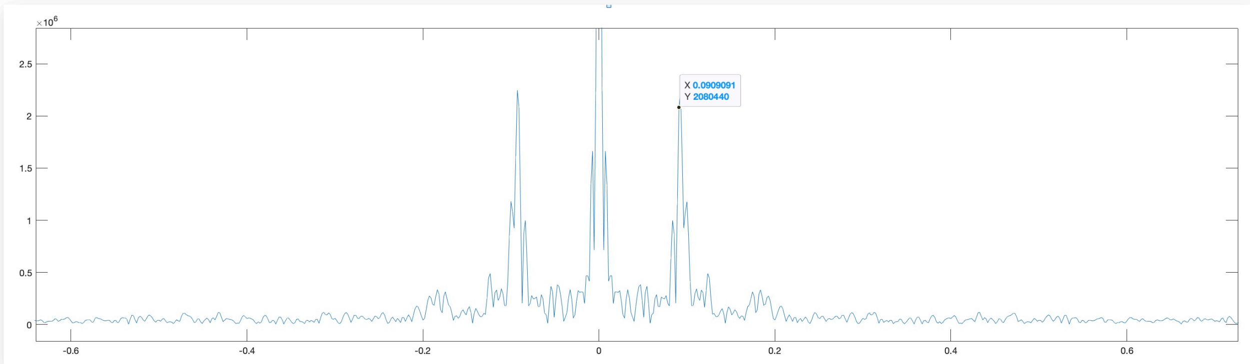
$$\mathbb{E}(X \mid (Y,Z) = (y,z)) = \int_{-\infty}^{\infty} x f_{X|(Y,Z)}(x \mid (y,z)\,\mathrm{d}x = \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\,\mathrm{d}x = \mathbb{E}(X \mid Y = y).$$

□

Problem 3

...

Solution. The plot does not overlap completely with what is suggested by the hint; nevertheless there is a spike (and two minor ones) at around $x = 1/11$ (and at around $x = -1/11$ as well).



Problem 4

Let $\theta \in \mathbb{R}$ be an unknown parameter. Consider the density

$$f_\theta(x) := \begin{cases} \exp(-(x-\theta)) & x \geqslant \theta \\ 0 & x < \theta. \end{cases}$$

Suppose $X_1, \ldots, X_n$ is a random sample of size n such that each X_i has density f_θ. Show that $X_{(1)} = \min\limits_{1 \leqslant i \leqslant n} X_i$ is a sufficient statistic for θ.

Proof. We first write $f_\theta(x) = \exp(-(x-\theta))\chi_{[\theta,\infty)}(x_i)$. Since X_i's are i.i.d., for $\vec{x} := (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f_\theta(\vec{x}) = \prod_{i=1}^{n} f_\theta(x_i) = \exp(n\theta - \sum_{i=1}^{n} x_i) \prod_{i=1}^{n} \chi_{[\theta,\infty)}(x_i).$$

Note that $f_\theta(\vec{x}) \neq 0$ if and only if $x_i \geqslant \theta$ for all i, i.e., $X_{(1)} \geqslant \theta$. That is,

$$f_\theta(\tilde{x}) = \exp(n\theta - \sum_{i=1}^{n} x_i)\chi_{[\theta,\infty)}(X_{(1)}).$$

Therefore, $f_\theta(\tilde{x})$ admits a factorization

$$f_\theta(\tilde{x}) = \underbrace{\exp(n\theta)\chi_{[\theta,\infty)}(X_{(1)})}_{g_\theta(X_{(1)})} \cdot \underbrace{\exp(-\sum_{i=1}^{n} x_i)}_{h(x)},$$

which by the factorization theorem shows $X_{(1)}$ is sufficient. □