MATH 281A: Solution to Homework #1

Xiao Pu with special thanks to Jongha Ryu

October 10, 2016

Problem 1. *(Problem 1.9.1.2).*

Solution. Since $X_1, \ldots, X_n$ are uncorrelated $\sim (\theta, \sigma^2)$, we have $\mathsf{E}[X_i^2] = \mathrm{Var}(X_i) + \mathsf{E}[X_i]^2 = \sigma^2 + \theta^2$ for all i and $\mathsf{E}[X_i X_j] = \theta^2$ for all $i \neq j$. Let $Y = \sum_i \alpha_i X_i$ where $\sum_i \alpha_i = 1$. Then clearly $\mathsf{E}[Y] = \theta$, and

$$\begin{aligned}
\mathsf{E}[Y^2] &= \sum_i \alpha_i^2 \, \mathsf{E}[X_i^2] + \sum_{i \neq j} \alpha_i \alpha_j \, \mathsf{E}[X_i X_j] \\
&= \theta^2 + \sigma^2 \sum_i \alpha_i^2 \\
&\geq \theta^2 + \sigma^2 \frac{1}{n} \Big(\sum_i \alpha_i \Big)^2 = \theta^2 + \frac{\sigma^2}{n},
\end{aligned}$$

where Cauchy-Schwarz inequality was used in the last inequality, so with equality if and only if $\alpha_1 = \ldots = \alpha_n = \frac{1}{n}$. Thus,

$$\mathrm{Var}(Y) = \mathsf{E}[Y^2] - \mathsf{E}[Y]^2 \geq \frac{\sigma^2}{n}.$$

Thus, the mean has the smallest variance. □

Problem 2. *(Problem 1.9.1.3).*

Solution of (a). We have $\mathsf{E}[X_i^2] = \frac{\sigma^2}{\beta_i} + \theta^2$ for each i and $\mathsf{E}[X_i X_j] = \theta^2$ for all $i \neq j$. Thus, $\mathsf{E}[Y_i] = \theta$ for all i, and

$$\begin{aligned}
\mathsf{E}[Y^2] &= \sum_i \alpha_i^2 \, \mathsf{E}[X_i^2] + \sum_{i \neq j} \alpha_i \alpha_j \, \mathsf{E}[X_i X_j] \\
&= \theta^2 + \sigma^2 \sum_i \frac{\alpha_i^2}{\beta_i} \\
&\geq \theta^2 + \sigma^2 \frac{\left(\sum_i \alpha_i \right)^2}{\sum_i \beta_i} = \theta^2 + \frac{\sigma^2}{\sum_i \beta_i},
\end{aligned}$$

with equality if and only if $\frac{\beta_1}{\alpha_1} = \ldots = \frac{\beta_n}{\alpha_n} = \text{const.}$, which implies for each i, $\alpha_i = \beta_i / \sum_i \beta_i$. Thus,

$$\mathrm{Var}(Y) = \mathsf{E}[Y^2] - \mathsf{E}[Y]^2 \geq \frac{\sigma^2}{\sum_{i=1}^n \beta_i},$$

and the minimum is achieved by the estimator $Y = \sum_i \frac{\beta_i}{\sum_j \beta_j} X_i$. □

Solution of (b). From the conditions given, $\mathsf{E}[X_i] = \theta^2 + \sigma^2$ for all i, and $\mathsf{E}[X_i X_j] = \mathrm{Cov}(X_i, X_j) + \mathsf{E}[X_i]\,\mathsf{E}[X_j] = \theta^2 + \rho\sigma^2$. Thus,

$$\begin{aligned}
\mathsf{E}[Y^2] &= \sum_i \alpha_i^2\, \mathsf{E}[X_i^2] + \sum_{i \neq j} \alpha_i \alpha_j\, \mathsf{E}[X_i X_j] \\
&= \sum_i \alpha_i^2 (\sigma^2 + \theta^2) + \sum_{i \neq j} \alpha_i \alpha_j (\theta^2 + \rho\sigma^2) \\
&= \theta^2 + \rho\sigma^2 + \sigma^2 (1 - \rho) \sum_i \alpha_i^2 \\
&\geq \theta^2 + \sigma^2 \Big(\frac{1}{n} + \frac{n-1}{n} \rho \Big).
\end{aligned}$$

Like in the previous problems, we use Cauchy-Schwarz inequality, and we have equality achieved if and only if $\alpha_1 = \ldots = \alpha_n = \frac{1}{n}$. Finally we have

$$\mathrm{Var}(Y) = \mathsf{E}[Y^2] - \mathsf{E}[Y]^2 \geq \sigma^2 \Big(\frac{1}{n} + \frac{n-1}{n} \rho \Big),$$

and this minimum can be only achieved by the mean $\overline{X}$. □

Problem 3. *(Problem 1.9.1.4).* Assuming $X \sim (\theta, \sigma^2)$, $Y \sim (\theta, \tau^2)$, and $\mathrm{corr}(X, Y) = \rho$, we have $\mathsf{E}[XY] = \rho\sigma\tau + \theta^2$.

Solution of (a).

$$\begin{aligned}
\mathrm{Var}\Big(\frac{X+Y}{2} \Big) &= \mathsf{E}\Big[\Big(\frac{X+Y}{2} \Big)^2 \Big] - \Big(\mathsf{E}\Big[\frac{X+Y}{2} \Big] \Big)^2 \\
&= \frac{1}{4}\, \mathsf{E}[X^2 + Y^2] + \frac{1}{2}\, \mathsf{E}[XY] - \theta^2 \\
&= \frac{1}{4} (2\theta^2 + \sigma^2 + \tau^2) + \frac{1}{2} (\rho\sigma\tau - \theta^2) \\
&= \frac{1}{4} (\sigma^2 + \tau^2 + 2\rho\sigma\tau) \\
&> \sigma^2 = \mathrm{Var}(X)
\end{aligned}$$

is the equivalent condition of $\mathrm{Var}(X) < \mathrm{Var}((X+Y)/2)$. □

Solution of (b).

$$\begin{aligned} f(\alpha) &= \mathrm{Var}(\alpha X + (1-\alpha)Y) \\ &= \mathsf{E}[(\alpha X + (1-\alpha)Y)^2] - (\mathsf{E}[\alpha X + (1-\alpha)Y])^2 \\ &= \alpha^2(\theta^2+\sigma^2) + (1-\alpha)^2(\theta^2+\tau^2) + 2\alpha(1-\alpha)(\rho\sigma\tau+\theta^2) - \theta^2 \\ &= \alpha^2(\sigma^2+\tau^2-2\rho\sigma\tau) + 2\alpha(-\tau^2+\rho\sigma\tau) + \tau^2. \end{aligned}$$

Thus, the minimizer of $f(\alpha)$ is

$$\alpha_* = \frac{\tau(\tau-\rho\sigma)}{\sigma^2+\tau^2-2\rho\sigma\tau}.$$

To make this minimizer negative, we require the equivalent condition $\tau < \rho\sigma$. □

Problem 4. *(Problem 1.9.1.7).*

Solution of (a). $\mathsf{P}(X \le m) = 1 - \mathsf{P}(X > m) \ge \frac{1}{2} \Leftrightarrow \mathsf{P}(X > m) \le \frac{1}{2}$, and $\mathsf{P}(X \ge m) = 1 - \mathsf{P}(X < m) \ge \frac{1}{2} \Leftrightarrow \mathsf{P}(X < m) \le \frac{1}{2}$. □

Solution of (b). Let $m_0 = \inf\{m\colon \mathsf{P}(X \le m) \ge \frac{1}{2}\}$ and $m_1 = \sup\{m\colon \mathsf{P}(X \ge m) \ge \frac{1}{2}\}$. We now only need to prove that m_0 and m_1 are medians. By definition, for any $\epsilon > 0$, we have $\mathsf{P}(X \le m_0 - \epsilon) < \frac{1}{2}$, which implies $\mathsf{P}(X > m_0 - \epsilon) > \frac{1}{2}$. By taking $\epsilon \to 0$, $\mathsf{P}(X \ge m_0) \ge \frac{1}{2}$, and thus m_0 is a median. We can prove that m_1 is median by the same reasoning. □

Problem 5. *(Problem 1.9.1.8).*

Solution. Let $m \in [m_0, m_1]$ is a median. Note that $\phi(a) = \mathsf{E}\,|X-a| = \mathsf{E}[(X-a)(1-2\mathbf{1}_{\{X\le m\}})]$ for any a. Therefore, for $c > m_1$ we have

$$\begin{aligned} \mathsf{E}\,|X-c| - \mathsf{E}\,|X-m| &= \mathsf{E}[(X-c)(1-2\mathbf{1}_{\{X<c\}})] - \mathsf{E}[(X-m)(1-2\mathbf{1}_{\{X\le m\}})] \\ &= \mathsf{E}[(X-c)((1-2\mathbf{1}_{\{X<c\}}) - (1-2\mathbf{1}_{\{X\le m\}}))] - \mathsf{E}[(c-m)(\mathbf{1}_{\{X>m\}} - \mathbf{1}_{\{X\le m\}})] \\ &= 2\,\mathsf{E}[(\mathbf{1}_{\{X<c\}} - \mathbf{1}_{\{X\le m\}})(c-X)] + (c-m)\,\mathsf{E}[\mathbf{1}_{\{X\le m\}} - \mathbf{1}_{\{X>m\}} \\ &= (c-m)(\mathsf{P}(X\le m) - \mathsf{P}(X>m)) + 2\int_{m<x<c} (c-x)\,d\,\mathsf{P}(x) \\ &\ge 0, \end{aligned}$$

because $\mathsf{P}(X \le m) \ge \frac{1}{2}$, and $\mathsf{P}(X > m) \le \frac{1}{2}$, and this implies that $\mathsf{E}\,|X-c| \ge \mathsf{E}\,|X-m|$ for any $c > m_1$. We can prove that this holds for $c < m_0$ by the same reasoning. Hence, any median minimizes $\phi(a)$. □

Problem 6. *(Problem 1.9.1.9).*

Solution of (a). Obvious from definition. □

Solution of (b). If we assume uniform distribution over the finite set $\{x_1, \ldots, x_n\}$, then $\sum_i |x_i - a| = n\phi(a)$ using the notation in Problem 1.9.1.8. Hence, any median is a minimizer of the sum of absolute deviations. □

Solution of (c). Without loss of generality, we can assume that $x_i \neq 0$ for all i. (If $x_i = 0$, this would make the term $|y_i - bx_i| = |y_i|$, a constant.) Then,

$$\begin{aligned}\arg\min_b \sum_i |y_i - bx_i| &= \arg\min_b \sum_i \frac{|x_i|}{\sum_j |x_j|}\left|\frac{y_i}{x_i} - b\right| \\ &= \arg\min_b \mathsf{E}[|Z - b|],\end{aligned}$$

where we let Z be a random variable taking values from $\{y_1/x_1, \ldots, y_n/x_n\}$, and let pmf of Z be $(|x_1|, \ldots, |x_n|)/(\sum_j |x_j|)$. Therefore, a minimizer b is any median of the random variable Z. $\square$

Problem 7. *(Problem 1.9.1.10).*

Solution. Note that

$$\arg\min_a \sum_{i=1}^n (h(x_i) - h(a))^2 = \arg\min_a \Big(h(a)^2 - \frac{2}{n}h(a)\sum_{i=1}^n h(x_i)\Big) = h^{-1}\Big(\frac{1}{n}\sum_{i=1}^n h(x_i)\Big),$$

since $h(\cdot)$ is a monotone function. One can easily check that 1) $h(x) = x$ gives AM, 2) $h(x) = \log x$ gives GM, and 3) $h(x) = \frac{1}{x}$ gives HM. $\square$

Problem 8. *(Problem 1.9.1.12).*

Solution of (a). Let $f_k(x) = \frac{k-1}{2}\frac{1}{(1+|x|)^k}$. f_k is a density function because

$$\begin{aligned}\int_{-\infty}^{\infty} f_k(x)dx &= \frac{k-1}{2}\int_{-\infty}^{\infty}\frac{dx}{(1+|x|)^k} \\ &= (k-1)\int_0^{\infty}(1+x)^{-k}dx \\ &= (k-1)\left(-\frac{(1+x)^{-(k-1)}}{k-1}\right)_0^{\infty} = 1.\end{aligned}$$

Now assume $1 \leq l < k-1$. For even l,

$$\begin{aligned}\mathsf{E}[X^l] &= \int_{-\infty}^{\infty} x^l f_k(x)dx \\ &= \int_{-\infty}^{\infty} x^l \frac{k-1}{2}\frac{dx}{(1+|x|)^k} \\ &= (k-1)\int_0^{\infty} x^l(1+x)^{-k}dx \\ &= (k-1)\left(-\frac{x^l(1+x)^{-(k-1)}}{k-1}\bigg|_0^{\infty} + \frac{l}{k-1}\int_0^{\infty} x^{l-1}(1+x)^{-(k-1)}dx\right) \\ &= (k-1)\frac{l}{k-1}\int_0^{\infty} x^{l-1}(1+x)^{-(k-1)}dx \\ &= \ldots\end{aligned}$$

$$= \frac{l!(k-l-1)!}{(k-2)!} \int_0^\infty (1+x)^{-(k-l)} dx$$
$$< \infty,$$

since $k - l > 1$. For odd l, since $\mathsf{E}[|X|^l] < \infty$ by the above result, we have $\mathsf{E}[X^l] = 0$ by symmetry. □

Solution of (b). Note that

$$2\int f^2(x)dx = \frac{(k-1)^2}{2}\int_{-\infty}^{\infty} \frac{dx}{(1+|x|)^{2k}} = (k-1)^2 \int_0^\infty (1+x)^{-2k} dx = \frac{(k-1)^2}{2k-1} < \frac{k-1}{2} = f(0),$$

because $2(k-1) < 2k-1$ for all k. □

Problem 9. *(Problem 1.9.4.1).*

Solution. Recall: A scale family of distributions has densities of the form $f(x;\sigma) = \frac{1}{\sigma}\psi(\frac{x}{\sigma})$ where $\sigma > 0$ and a location family of distributions has densities of the form $f(x;\mu) = \psi(x-\mu)$ where $-\infty < \mu < \infty$. Let $Y = \log X$, and let $f_X(x;\sigma) = \frac{1}{\sigma}\psi(\frac{x}{\sigma})$ be the pdf of X. Then we have $F_Y(y) = F_X(e^y)$, and thus $f_Y(y) = e^y f_X(e^y) = e^{y-\log\sigma}\psi(e^{y-\log\sigma}) = \varphi(y - \log\sigma)$ for some function $\varphi(\cdot)$. This proves that $Y = \log X$ is a location family. □

Problem 10. *(Problem 1.9.4.2).*

Solution. We want to prove that if $X \sim U(0,\theta)$, then $-\log X \sim \exp$. The cdf of X is $F_X(x) = \frac{x}{\theta}$ for $x \in [0,\theta]$. Let $Y = -\log X \in [-\log\theta, \infty)$. Then, $F_Y(y) = \mathsf{P}(-\log X \le y) = \mathsf{P}(X \ge e^{-y}) = 1 - F_X(e^{-y}) = 1 - e^{-y}/\theta$ for $y \in [-\log\theta, \infty)$. Hence, we get the pdf of Y $f_Y(y) = e^{-y}\theta$ for $y \in [-\log\theta, \infty)$, which implies that $Y \sim E(-\log\theta, 1)$. □

Problem 11. *(Problem 1.9.4.13(b)).*

Solution. We have $U \sim E(0,1)$, $X = bU^{1/c}$, where $b, c > 0$. Then the cdf of X becomes $F_X(x) = \mathsf{P}(bU^{1/c} \le x) = \mathsf{P}(Y \le (x/b)^c) = 1 - e^{-(x/b)^c}$, which implies that the pdf of X is $f_X(x) = \frac{c}{b}(\frac{x}{b})^{c-1}e^{-(\frac{x}{b})^c}$, for $x > 0$. □

Problem 12. *(Additional problem 1).* Suppose $X_1, \ldots, X_n \sim$ i.i.d. , and $\mathsf{E}\, X_1^4 < \infty$. Then prove $sk(\overline{X}) = sk(X_1)/\sqrt{n}$ and $ku(\overline{X}) = ku(X_1)/n$.

Solution of (a).

$$\mu_3 = \mathsf{E}[(\overline{X}-\mu)^3] = \mathsf{E}[\frac{1}{n}\sum_i (X_i - \mu)]^3$$
$$= \frac{1}{n^3}\,\mathsf{E}[\sum_i (X_i-\mu)]^3 = \frac{n}{n^3}\,\mathsf{E}(X_i-\mu)^3 = \frac{\mathsf{E}(X_i-\mu)^3}{n^2}$$

Therefore,

$$\begin{aligned} sk(\overline{X}) &= \frac{\mu_3}{(var(\overline{X}))^{3/2}} \\ &= \frac{\mathsf{E}(X_1-\mu)^3}{n^2\cdot(\frac{\mathsf{E}(X_1-\mu)^2}{n})^{3/2}} \\ &= \frac{1}{\sqrt{n}}\frac{\mathsf{E}(X_1-\mu)^3}{[\mathsf{E}(X_1-\mu)^2]^{3/2}} \\ &= \frac{1}{\sqrt{n}} sk(X_1). \end{aligned}$$

As $n \to \infty$, $sk(\overline{X}) \to 0$. □

Solution of (b).

$$\begin{aligned} \mu_4 = \mathsf{E}[(\overline{X}-\mu)^4] &= \mathsf{E}[\frac{1}{n}\sum_i (X_i-\mu)]^4 \\ &= \frac{1}{n^4}\{\sum_i \mathsf{E}(X_i-\mu)^4 + \sum_i \sum_{i<j\le n} 6\,\mathsf{E}(X_i-\mu)]^2\,\mathsf{E}(X_j-\mu)]^2\} \\ &= \frac{1}{n^4}\{n\,\mathsf{E}(X_1-\mu)^4 + 3n(n-1)[\mathsf{E}(X_1-\mu)^2]^2\} \end{aligned}$$

Therefore,

$$\begin{aligned} ku(\overline{X}) &= \frac{\mathsf{E}[(\overline{X}-\mu)^4]}{var(\overline{X})^2} - 3 \\ &= \frac{\frac{1}{n^4}\{n\,\mathsf{E}(X_1-\mu)^4 + 3n(n-1)[\mathsf{E}(X_1-\mu)^2]^2\}}{\frac{[var(X_1)]^2}{n^2}} - 3 \\ &= \frac{1}{n}\{\frac{\mathsf{E}(X_1-\mu)^4}{[var(X_1)]^2} - 3\} \\ &= \frac{1}{n} ku(X_1). \end{aligned}$$

As $n \to \infty$, $ku(\overline{X}) \to 0$. □

Problem 13. *(Additional problem 2).* Show that the usually-used estimator of σ^2 is always unbiased.

Solution of (13). We assume $X_1, \ldots, X_n \sim$ i.i.d. (μ, σ^2). Then for all i $\mathsf{E}[X_i^2] = \mu^2 + \sigma^2$ and $\mathsf{E}[\overline{X}^2] = \mu^2 + \frac{\sigma^2}{n}$. Now we have, for each i,

$$\mathsf{E}(X_i - \overline{X})^2 = \mathsf{E}\,X_i^2 + \mathsf{E}\,\overline{X}^2 - \mathsf{E}\,X_i\overline{X}$$

$$= (\mu^2 + \sigma^2) + \Big(\mu^2 + \frac{\sigma^2}{n}\Big) - \frac{2}{n}\big((\mu^2 + \sigma^2) + (n-1)\mu^2\big)$$
$$= \big(1 - \frac{1}{n}\big)\sigma^2.$$

Hence, by linearity of expectation,

$$\mathsf{E}\Big[\sum_{i=1}^{n}(X_i - \overline{X})^2\Big] = (n-1)\sigma^2. \qquad \square$$

Problem 14. *(Additional problem 3).* What is the variance of the sample variance (the estimator in problem 13)?

Solution of (14). Google "variance of the sample variance", you'll find the best answer comes in here "http://math.stackexchange.com/questions/72975/variance-of-sample-variance". In the future, use Google to help yourself learn Statistics and whatever you want! $\square$