Solutions to Math 481 Review Problems

1. (a) We compute

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\mu)^2 &= \sum_{i=1}^{n}\big((X_i-\overline{X})+(\overline{X}-\mu)\big)^2\\ &= \sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{i=1}^{n}2(X_i-\overline{X})(\overline{X}-\mu)+\sum_{i=1}^{n}(\overline{X}-\mu)^2.\end{aligned}$$

Now

$$\begin{aligned}\sum_{i=1}^{n}2(X_i-\overline{X})(\overline{X}-\mu) &= 2(\overline{X}-\mu)\cdot(\sum_{i=1}^{n}X_i-\sum_{i=1}^{n}\overline{X})\\ &= 2(\overline{X}-\mu)\cdot(n\overline{X}-n\overline{X})=0,\end{aligned}$$

and the result follows.

(b) We compute

$$\begin{aligned}E[S^2] &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]\\ &= E\left[\frac{1}{n-1}\left(\left(\sum_{i=1}^{n}(X_i-\mu)^2\right)-n(\overline{X}-\mu)^2\right)\right] \quad \text{(by (a))}\\ &= \frac{1}{n-1}\left(\left(\sum_{i=1}^{n}E[(X_i-\mu)^2]\right)-nE[(\overline{X}-\mu)^2]\right) \quad \text{(by linearity of } E\text{)}\\ &= \frac{1}{n-1}(n\sigma^2-\sigma^2) \quad (\text{since } \mathrm{Var}(\overline{X})=\sigma^2/n)\\ &= \sigma^2.\end{aligned}$$

2. (a) Independence implies that the moment generating function for $Z = X_1+\cdots+X_n$ satisfies

$$M_Z(t)=\prod_{i=1}^{n}M_{X_i}(t)=\prod_{i=1}^{n}(1-2t)^{-1/2}=(1-2t)^{-n/2}.$$

Thus Z has the same moment generating function as does a chi-square random variable with n degrees of freedom, so these two random variables must be the same.

(b) Since $M_{cY}(t) = M_Y(ct) = (1 - \beta ct)^{-\alpha}$, as above we conclude that cY has gamma distribution with parameters α and $c\beta$.

3. (a) μ_A will be larger than μ_B, since the inequality used to prove Chebyshev's result can be improved by using the actual distribution.

(b) Chebychev's inequality says that $P(|X - \mu_A| \geq c\sigma) \leq 1/c^2$, where the random variable X is the amount of soda in a can (hence $\sigma = 0.1$). Putting $1/c^2 = 0.05$, we have $P(|X - \mu_A| < 0.1\sqrt{20}) \geq 0.95$. Thus $P(X > \mu_A - 0.1\sqrt{20}) \geq 0.95$, so the order's requirement is satisfied if $\mu_A - 0.1\sqrt{20} = 12$, or in other words $\mu_A = 12 + 1/\sqrt{5}$, approximately 12.447. On the other hand, if X is a normal random variable with mean μ_B and standard deviation 0.1, then $Z = (X - \mu_B)/0.1$ has standard normal distribution, so

$$P(X > 12) = P(Z > 10(12 - \mu_B)).$$

From the tables, $P(Z > -1.645) = 0.95$, so $10(12 - \mu_B) = -1.645$ and thus $\mu_B = 12.1645$. And indeed, as we decided in (a) before doing any computations, we find that $\mu_A > \mu_B$.

(c) This time let $X_1, \ldots, X_n$ (with $n = 10,000$) be the amount of soda put in each can. These are independent random variables with mean 12.1 and standard deviation 0.1. Engineer B assumes they are normally distributed, so $\overline{X}$ has normal distribution with mean 12.1 and standard deviation $0.1/\sqrt{n} = 1/1000$. Thus $Z = 1000(\overline{X} - 12.1)$ has standard normal distribution. He then computes the required amount of bulk soft drink S_B as follows:

$$0.99 < P(n\overline{X} < S_B) = P(Z < \frac{S_B}{10} - 12100),$$

and from the tables we see that $S_B/10 - 12100 = 2.33$, so $S_B = 121023.3$. The central limit theorem says that the distribution of $\overline{X}$ is approximately normal, even if the X_i are not normally distributed. Thus it gives the exact same answer, so $S_A = S_B$.

4. (a) We know that $Z = (\overline{X} - \mu)/(\sigma/\sqrt{n}) = (\overline{X} - \mu)/\sqrt{5/3}$ has the standard normal distribution. Thus, $P(|\overline{X} - \mu| < 3) = P(|Z| < 3\sqrt{3/5}) = 2P(0 < Z < 3\sqrt{3/5})$, so (from the table) the probability is approximately $2(0.4898) = 0.9796$.

(b) We know that $X = \frac{n-1}{\sigma^2}S^2$ has the chi-square distribution with $n - 1$ degrees of freedom. Here $n = 9$, so $P(S^2 > a\sigma^2) = P(X > 8a)$. From the table, $P(X > 8a) < 0.01$ when $8a = 20.09$, so $a = 2.51$.

(c) We know that $T = (\overline{X} - \mu)/(S/\sqrt{n})$ has the t-distribution with $n - 1$ degrees of freedom. Here $n = 9$, so $P(|\overline{X} - \mu| > b) = P(|T| > b\sqrt{3/5})$. From the table, $P(T > b\sqrt{3/5}) = (0.05)/2$ when $b\sqrt{3/5} = 2.306$, so $b = 2.977$.

5. (a) The probability density function for Y_3 is

$$f_{Y_3}(y) = 12 \cdot \left(\int_0^y \frac{1}{\beta} dx \right)^2 \cdot \frac{1}{\beta} \cdot \left(\int_y^\beta \frac{1}{\beta} dx \right)$$
$$= \frac{12y^2(\beta - y)}{\beta^4}$$

for $0 < y < \beta$ (and 0 otherwise). Thus

$$\begin{aligned} E[Y_3] &= \int_0^\beta y f_{Y_3}(y) dy \\ &= \int_0^\beta \frac{12}{\beta^4} y^3 (\beta - y) dy \\ &= \frac{12}{\beta^4} \left(\frac{\beta}{4} y^4 - \frac{1}{5} y^5 \right) \Big|_0^\beta \\ &= \frac{3\beta}{5}. \end{aligned}$$

(b) Since E is linear, we can take $c = 5/3$.

(c) We compute

$$\begin{aligned} E[Y_3^2] &= \int_0^\beta y^2 f_{Y_3}(y) dy \\ &= \int_0^\beta \frac{12}{\beta^4} y^4 (\beta - y) dy \\ &= \frac{12}{\beta^4} \left(\frac{\beta}{5} y^5 - \frac{1}{6} y^6 \right) \Big|_0^\beta \\ &= \frac{2\beta^2}{5}, \end{aligned}$$

so $\mathrm{Var}(Y_3) = E[Y_3^2] - E[Y_3]^2 = \beta^2/25$ and thus $\mathrm{Var}(cY_3) = \beta^2/9$. Likewise, $f_{Y_4}(y) = 4y^3/\beta^4$ if $0 < y < \beta$, so $E[Y_4] = 4\beta/5$ and $E[Y_4^2] = 2\beta^2/3$, whence $\mathrm{Var}(Y_4) = 2\beta^2/75$ and $\mathrm{Var}(5Y_4/4) = \beta^2/24$. Thus the relative efficiency is $\mathrm{Var}(cY_3)/\mathrm{Var}(5Y_4/4) = 8/3$.

6. Letting $\overline{X}$ be the sample mean, and μ and σ^2 the population mean and population variance, we know that $T = (\overline{X} - \mu)/(S/\sqrt{n})$ has the t-distribution with $n-1$ degrees of freedom. We want $P(|\overline{X} - \mu| > 3S) \leq 0.01$, or equivalently $P(|T| > 3\sqrt{n}) \leq 0.01$; since the t-distribution is symmetric around $t = 0$, we can rewrite this as $P(T > 3\sqrt{n}) \leq 0.005$. For fixed n, let $t_{0.005,n-1}$ satisfy $P(T > t_{0.005,n-1}) = 0.005$. We need an n for which $3\sqrt{n} \geq t_{0.005,n-1}$; since $t_{0.005,3} < 6$, we can take $n = 4$.

7. The probability density function of $Y = Y_{n-1}$ is

$$f_Y(y) = n(n-1) \left(\int_0^y \frac{1}{\beta} dx \right)^{n-2} \cdot \frac{1}{\beta} \cdot \int_y^\beta \frac{1}{\beta} dx = \frac{n(n-1)}{\beta^n} y^{n-2} (\beta - y)$$

if $0 < y < \beta$ (and $f_Y(y) = 0$ otherwise). Thus,

$$\begin{aligned} E[Y] &= \int_0^\beta y f_Y(y) dy \\ &= \frac{n(n-1)}{\beta^n} \int_0^\beta y^{n-1}(\beta - y) dy \\ &= \frac{n(n-1)}{\beta^n} \left(\frac{y^n}{n}\beta - \frac{y^{n+1}}{n+1} \right) \bigg|_0^\beta \\ &= \frac{n-1}{n+1}\beta. \end{aligned}$$

Thus Y is biased, since $E[Y] \neq \beta$; but $E[Y] \to \beta$ as $n \to \infty$, so Y is asymptotically unbiased. In view of this, to show Y is consistent it will suffice to prove that $\mathrm{Var}(Y) \to 0$ as $n \to \infty$. To this end we compute

$$\begin{aligned} E[Y^2] &= \int_0^\beta y^2 f_Y(y) dy \\ &= \frac{n(n-1)}{\beta^n} \int_0^\beta y^n(\beta - y) dy \\ &= \frac{n(n-1)}{\beta^n} \left(\frac{y^{n+1}}{n+1}\beta - \frac{y^{n+2}}{n+2} \right) \bigg|_0^\beta \\ &= \frac{n(n-1)}{(n+1)(n+2)}\beta^2, \end{aligned}$$

so

$$\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 = \left(\frac{n(n-1)}{(n+1)(n+2)} - \frac{(n-1)^2}{(n+1)^2} \right) \beta^2 = \frac{2(n-1)\beta^2}{(n+1)^2(n+2)},$$

and indeed $\mathrm{Var}(Y) \to 0$ as $n \to \infty$.

8. The new program should be written so that, each time it is called, it makes 8 calls to the first program, yielding values $x_1, \ldots, x_8$, and the output of the new program is $y = x_1^2 + x_2^2 + \cdots + x_8^2$.

9. Let $X_1, \ldots, X_{50}$ be a random sample from a population having chi-square distribution with $\nu = 1$ degree of freedom. Each of these has mean $\mu = 1$ and standard deviation $\sigma = \sqrt{2}$. The central limit theorem says that the distribution of $Z = (\overline{X} - \mu)/(\sigma/\sqrt{50}) = 5(\overline{X} - 1)$ is approximately the standard normal distribution. From the tables, $P(Z \geq 2.33) \approx 0.01$ and $P(Z \geq -2.33) \approx 0.99$, so $P(50\overline{X} \geq 73.3) \approx 0.01$ and $P(50\overline{X} \geq 26.7) \approx 0.99$. Since $50\overline{X}$ has chi-square distribution with 50 degrees of freedom (by problem 2), this implies $\chi^2_{0.01,50} \approx 73.3$ and $\chi^2_{0.99,50} \approx 26.7$.

10. The first equation is true in general, as it simply says that E is linear. The second equation is true if the random variables $X_1, \ldots, X_n$ are independent – if

the variables were not known to be independent, the equation would be

$$\operatorname{Var}\left(\sum_{i-1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{Cov}(X_i, X_j).$$

Here is a simple counterexample to the second equation in this problem. Take $n = 2$ and $a_1 = a_2 = 1$ and $X_2 = X_1$. Then the second equation says that $\operatorname{Var}(X_1 + X_2) = \operatorname{Var}(X_1) + \operatorname{Var}(X_1)$, or equivalently $\operatorname{Var}(2X_1) = 2\operatorname{Var}(X_1)$, which is false whenever $\operatorname{Var}(X_1) \neq 0$.

11. Let $X_1, \ldots, X_n$ be independent random variables all of which have the same probability density function $f(x; \theta)$, which depends on a parameter θ. Let Y be some function of $X_1, \ldots, X_n$ – so Y is a *statistic*, or equivalently Y is an *estimator* of θ. Then Y is a *sufficient* statistic if the conditional joint probability density function $f(X_1, \ldots, X_n | Y; \theta)$ does not depend on the value of θ.

For example, let $f(x; \theta)$ be the uniform distribution on the interval $[0, \theta]$, and let $Y = \max(X_1, \ldots, X_n)$ be the n^{th} order statistic. Then the joint probability density function for $X_1, \ldots, X_n$ is

$$f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n},$$

so long as $0 < x_1, \ldots, x_n < \theta$ (otherwise $f = 0$). Thus

$$f(x_1, \ldots, x_n; \theta) = g(y, \theta) \cdot h(x_1, \ldots, x_n),$$

where

$$g(y, \theta) = \begin{cases} \frac{1}{\theta^n} & \text{if } y < \theta \\ 0 & \text{otherwise} \end{cases}$$

and

$$h(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } 0 < x_1, \ldots, x_n \\ 0 & \text{otherwise,} \end{cases}$$

so the Fisher-Neyman Factorization Theorem implies that Y is a sufficient estimator of θ.