Econometrics HW4

Michael B. Nattinger*

September 30, 2020

1 Question 1

1.1 Show that $\bar{X}_{n+1} = (n\bar{X}_n + X_{n+1})/(n+1)$.

$$\begin{aligned}\bar{X}_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1} X_i = \frac{1}{n+1}\left(\left(\sum_{i=1}^{n} X_i\right) + X_{n+1}\right) \\ &= \frac{1}{n+1}\left(n\bar{X}_n + X_{n+1}\right).\end{aligned}$$

1.2 Show that $s^2_{n+1} = ((n-1)s^2_n + (n/(n+1))(X_{n+1} - \bar{X}_n)^2)/n$.

$$\begin{aligned}s^2_{n+1} &= \frac{1}{n}\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2 = \frac{1}{n}\sum_{i=1}^{n+1}(X_i - \bar{X}_n + \bar{X}_n + \bar{X}_{n+1})^2 \\ &= \frac{1}{n}\left((n-1)\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2 + n(\bar{X}_n - \bar{X}_{n+1}) + 2(\sum_{i=1}^{n}(X_j - \bar{X}_n)(\bar{X}_n - \bar{X}_{n+1})) + (X_{n+1} - \bar{X}_{n+1})\right) \\ &= \frac{1}{n}\left((n-1)\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2 + n(\bar{X}_n - \bar{X}_{n+1})^2 + (X_{n+1} - \bar{X}_{n+1})^2\right) \\ &= \frac{1}{n}\left((n-1)s^2_n + n\bar{X}^2_n - 2n\bar{X}_n\bar{X}_{n+1} + n\bar{X}^2_{n+1} + X^2_{n+1} - 2X_{n+1}\bar{X}_{n+1} + \bar{X}^2_{n+1}\right) \\ &= \frac{1}{n}\left((n-1)s^2_n + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2\right).\end{aligned}$$

2 Question 2

Define $\hat{\mu}_k := \frac{1}{n}\sum_{i=1}^{n} X_i^k$. We will show that this is unbiased.

$$\begin{aligned}E[\hat{\mu}_k] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^k\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i^k] \\ &= \frac{1}{n}\sum_{i=1}^{n}\mu_k \\ &= \mu_k.\end{aligned}$$

Thus, $\hat{\mu}_k$ is an unbiased estimator for μ_k.

*I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, and Danny Edgel.

3 Question 3

Define $\hat{m}_k := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$. This estimator is biased, which can be seen from the fact that $\hat{m}_2 = \hat{\sigma}^2 \neq s_n^2$, so $\hat{m}_2$ is not an unbiased estimator for $m_2 = \sigma^2$. There exists no general formula for an unbiased estimator of $m_k, k > 3$ to the best of my knowledge.

4 Question 4

$$\begin{aligned} Var(\hat{\mu}_k) &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i^k) = \frac{1}{n}(E[X_i^{2k}] - E[X_i^k]^2) \\ &= \frac{1}{n}(\mu_{2k} - \mu_k^2). \end{aligned}$$

5 Question 5

Note that $f(x) = x^2$ is convex. By Jensen's inequality,

$$E[s_n]^2 \leq E[s_n^2] = \sigma^2,$$

so $E[s_n] \leq \sigma$ as both are nonnegative.

6 Question 6

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

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

7 Question 7

$$
\begin{aligned}
Cov(\hat{\sigma}^2, \bar{X}_n) &= E[\hat{\sigma}^2(\bar{X}_n - E\bar{X}_n)] \\
&= E\left[\left(\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2 - (\bar{X}_n-\mu)^2\right)(\bar{X}_n-\mu)\right] \\
&= \frac{1}{n^2}E\left[\left(\sum_{i=1}^n (X_i-\mu)^2\right)\left(\sum_{i=1}^n X_i - \mu\right)\right] - E[(\bar{X}_n-\mu)^3] \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n E\left[(X_i-\mu)^3\right]\right) + \frac{1}{n^2}\left(\sum_{1\le i<j\le n}\left[(X_i-\mu)^2(X_j-\mu)\right]\right) - E[(\bar{X}_n-\mu)^3] \\
&= \frac{1}{n}E\left[(X_i-\mu)^3\right] - E[(\bar{X}_n-\mu)^3] \\
&= \frac{1}{n}E\left[(X_i-\mu)^3\right] \\
&- \frac{1}{n^3}\left(\sum_{i=1}^n E[(X_i-\mu)^3] + 3\sum_{i\neq j} E[(X_i-\mu)^2(X_j-\mu)] + 3\sum_{i\neq j\neq k} E[(X_i-\mu)(X_j-\mu)(X_k-\mu)]\right) \\
&= \frac{1}{n}E\left[(X_i-\mu)^3\right] - \frac{1}{n^2}E[(X_i-\mu)^3] \\
&= \left(\frac{1}{n} - \frac{1}{n^2}\right)E[(X_i-\mu)^3]
\end{aligned}
$$

This quantity will be 0 when X_i has no skewness.

8 Question 8

8.1 Find $E[\bar{X}_n]$

$$
\begin{aligned}
E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] &= \frac{1}{n}\sum_{i=1}^n E\left[X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mu_i.
\end{aligned}
$$

8.2 Find $Var(\bar{X}_n)$.

$$
\begin{aligned}
Var(\bar{X}_n) = \frac{1}{n^2}Var(\sum_{i=1}^n X_i) &= \frac{1}{n^2}\sum_{i=1}^n Var(X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^n \sigma_i^2.
\end{aligned}
$$

9 Question 9

$$\begin{aligned}
E[Q] = E\left[\sum_{i=1}^{r} X_i^2\right] &= \sum_{i=1}^{n} E\left[X_i^2\right] \\
&= \sum_{i=1}^{r} \mu_X^2 + \sigma_X^2 = \sum_{i=1}^{n} 0^2 + 1^2 \\
&= r.
\end{aligned}$$

$$\begin{aligned}
Var(Q) = E[Q^2] - E[Q]^2 &= E\left[\left(\sum_{i=1}^{r} X_i^2\right)^2\right] - r^2 \\
&= \sum_{i=1}^{r} E\left[X_i^4\right] + 2 \sum_{1\le i<j\le r} E\left[X_i^2 X_j^2\right] - r^2 \\
&= \sum_{i=1}^{r} 3 + 2 \sum_{1\le i<j\le r} (1)(1) - r^2 \\
&= 3r + 2\left(\frac{r!}{2(r-2)!}\right) - r^2 = 3r + r(r-1) - r^2 \\
&= 2r.
\end{aligned}$$

Note: we calculated $E[X^4] = 3$ for $X \sim N(0,1)$ in the previous problem set.

10 Question 10

We will first show that a sum of independent normal random variables is normal. Let $Z_i \sim N(\mu_i, \sigma_i^2) \forall i \in [1, \dots, n]$ for some $n \in \mathbb{N}$. Then the MGF of Z_i is $M_{Z_i}(t) = exp(\mu_i t + \frac{\sigma_i^2 t^2}{2})$.[1] Then,[2]

$$M_{\sum_{i=1}^n Z_i}(t) = \prod_{i=1}^{n} M_{Z_i}(t) = exp\left(\sum_{i=1}^{n}\left(\mu_i t + \frac{\sigma_i^2 t}{2}\right)\right) = exp\left(\left(\sum_{i=1}^{n}\mu_i\right)t + \left(\sum_{i=1}^{n}\sigma_i^2\right)\frac{t^2}{2}\right).$$

So $\sum_{i=1}^n Z_i$ is of the form of a normal random variable with mean $\sum_{i=1}^n \mu_i$ and variance $\sum_{i=1}^n \sigma_i^2$.

10.1 Find $E[\bar{X}_n - \bar{Y}_n]$

From the above, $\bar{X}_n, \bar{Y}_n$ are of the form of normal variables with means μ_X, μ_Y and variances $\frac{1}{n_1}\sigma_X^2, \frac{1}{n_2}\sigma_Y^2$. Then, $\bar{X}_n - \bar{Y}_n$ is also normal with mean $\mu_X - \mu_Y$.

10.2 Find $Var[\bar{X}_n - \bar{Y}_n]$

From the above, $\bar{X}_n - \bar{Y}_n$ is normal with variance $\frac{1}{n_1}\sigma_X^2 + \frac{1}{n_2}\sigma_Y^2$.

10.3 Find the distribution of $\bar{X}_n - \bar{Y}_n$

From the above, $\bar{X}_n - \bar{Y}_n$ is of the form of a normal random variable with mean $\mu_X - \mu_Y$ and variance $\frac{1}{n_1}\sigma_X^2 + \frac{1}{n_2}\sigma_Y^2$.

---

[1] $M_{Z_i}(t) = \int_{-\infty}^{\infty} exp(xt)\frac{1}{\sqrt{2\pi\sigma_i^2}} exp(-(x-\mu_i)^2/(2\sigma_i^2))dx$. Define $z = (x-\mu_i)/\sigma_i$. Then $M_{Z_i}(t) = exp(\mu_i t)\int_{-\infty}^{\infty} exp(z\sigma t)\frac{1}{\sqrt{2\pi\sigma_i^2}} exp(-z^2/2)(dx/dz)dz = exp(\mu_i t)\int_{-\infty}^{\infty} exp(z\sigma t)\frac{1}{\sqrt{2\pi}} exp(-z^2/2)dz = exp(\mu_i t)exp(\sigma_i^2 t^2/2) = exp(\mu_i t + \frac{\sigma_i^2 t^2}{2})$.

[2] $M_{\sum_{i=1}^n Z_i}(t) = E[exp(\sum_{i=1}^n tZ_i)] = E[\prod_{i=1}^n exp(tZ_i)] = \prod_{i=1}^n E[exp(tZ_i)] = \prod_{i=1}^n M_{Z_i}(t)$