# STAT 563 HW #6

*Maggie Buffum*

*May 30, 2019*

## Problem 10.19

Another way in which underlying assumptions can be violated is if there is correlation in the sampling, which can seriously affect the properties of the sample mean. Suppose we introduce correlation in the case discussed in Exercise 10.2.1; that is, we observe $X_1, \ldots, X_n$, where $X_i \sim n(\theta, \sigma^2)$, but the $X_i$s are no longer independent.

(a) For the equicorrelated case, that is, $Corr(X_i, X_j) = \rho$ for $i \neq j$, show that

$$Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{n-1}{n}\rho\sigma^2$$

so that $Var(\bar{X}) \not\to 0$ as $n \to \infty$.

The variance of $\bar{X}$ is

$$
\begin{aligned}
Var(\bar{X}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left(\sum_i Var(X_i) + 2\sum_{i>j} Cov(X_i, X_j)\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + 2\sum_{i>j} \rho_{i,j}\sigma_i\sigma_j\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + 2\sum_{i>j} \rho\sigma^2\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + 2\binom{n}{2}\rho\sigma^2\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + 2\frac{n!}{2!(n-2)!}\rho\sigma^2\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + n(n-1)\rho\sigma^2\right) \\
&= \frac{\sigma^2}{n} + \frac{n-1}{n}\rho\sigma^2
\end{aligned}
$$

(b) If the $X_i$s are observed through time (or distance), it is sometimes assumed the correlation decreases with time (or distance), with one specific model being $Corr(X_i, X_j) = \rho^{|i-j|}$. Show that in this case

$$Var(\bar{X}) = \frac{\sigma^2}{n} + \frac{2\sigma^2}{n^2}\frac{\rho}{1-\rho}\left(n + \frac{1-\rho^n}{1-\rho}\right)$$

so $Var(\bar{X}) \to 0$ as $n \to \infty$. (See Miscellanea 5.8.2 for another effect of correlation.)

---

From part (a), consider the variance of $\bar{X}$ written as

$$Var(\bar{X}) = \frac{1}{n^2}\left(n\sigma^2 + 2\sum_{i>j}\rho_{i,j}\sigma_i\sigma_j\right)$$

We just need to update the covariance term. Replacing with the new correlation coefficient, we have

$$Cov(X_i, X_j) = \sum_{i>j}\rho_{i,j}\sigma_i\sigma_j$$
$$= \sigma^2\sum_{i>j}\rho^{|i-j|}$$
$$= \sigma^2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\rho^{i-j}$$

Consider the first few cases: when $i = 2$, $\rho^{i-j} = \{\rho\}$. When $i = 3$, $\rho^{i-j} = \{\rho^2, \rho\}$. When $i = 4$, $\rho^{i-j} = \{\rho^3, \rho^2, \rho\}$. This continues until $i = n$ and we have the set $\rho^{i-j} = \{\rho^{n-1}, \rho^{n-2}, \ldots, \rho^3, \rho^2, \rho\}$. That is, $\rho$ occurs $n-1$ times, $\rho^2$ occurs $n-2$, etc., such that

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1}\rho^{i-j} = \sum_{i=1}^{n-1}(n-i)\rho^i = \frac{\rho}{1-\rho}\left(n - \frac{1-\rho^n}{1-\rho}\right)$$

(the known partial sum of a geometric series). Now we have that the variance of $\bar{X}$ is

$$Var(\bar{X}) = \frac{1}{n^2}\left(n\sigma^2 + 2\sum_{i>j}\rho_{i,j}\sigma_i\sigma_j\right)$$
$$= \frac{1}{n^2}\left(n\sigma^2 + 2\sigma^2\frac{\rho}{1-\rho}\left(n - \frac{1-\rho^n}{1-\rho}\right)\right)$$
$$= \frac{\sigma^2}{n} + \frac{2\sigma^2}{n^2}\frac{\rho}{1-\rho}\left(n - \frac{1-\rho^n}{1-\rho}\right)$$

---

(c) The correlation structure in part (b) arises in an *autoregressive AR(1) model*, where we assume that $X_{i+1} = \rho X_i + \delta_i$, with $\delta_i$ iid $n(0, 1)$. If $|\rho| < 1$ and we define $\sigma^2 = 1/(1-\rho^2)$, show that $Corr(X_1, X_i) = \rho^{i-1}$.

---

We know that
$$Corr(X_1, X_i) = \frac{Cov(X_1, X_i)}{\sqrt{\sigma_1^2\sigma_i^2}}$$

Since we are given that $\sigma_1^2 = \sigma_i^2 = \sigma^2$, we just need to find the covariance of $X_1$ and $X_i$, defined as $Cov(X_1, X_i) = E[X_1 X_i] - E[X_1]E[X_i]$. Note that each $X_i$ is conditional on $X_{i-1}$ such that using the law of iterative expectations, we have

$$E[X_i] = E[E(X_i|X_{i-1})] = E[E(\rho X_{i-1} + \delta_{i-1})] = \rho E[X_{i-1}]$$

Again using the law of iterative expectations, we find that

$$E[X_{i-1}] = \rho E[X_{i-2}]$$

The process continues until we have the expectation of $X_i$ in terms of $X_1$, that is,

$$EX_i = \rho E[X_{i-1}] = \rho^2 E[X_{i-2}] = \cdots = \rho^{i-1} E[X_1]$$

Next we need to find $E[X_1 X_i]$. Again using the law of iterative expectations we have that

$$E[X_1 X_i] = E[E(X_1 X_i | X_{i-1})] = E[E(X_1 | X_{i-1}) E(X_i | X_{i-1})] = E[X_1 \rho^{i-1} E[X_1]] = \rho^{i-1} E[X_1^2]$$

Finally we can evaluate the correlation of $X_1$ and $X_i$:

$$Corr(X_1, X_i) = \frac{Cov(X_1, X_i)}{\sqrt{\sigma^2 \sigma^2}} = \frac{\rho^{i-1} E[X_1^2] - \rho^{i-1}(E[X_1])^2}{\sigma^2} = \frac{\rho^{i-1} Var(X_1)}{\sigma^2} = \frac{\rho^{i-1} \sigma^2}{\sigma^2} = \rho^{i-1}$$