Solution Assignment 1: EE425 1/2018
APPENDIX B
Fundamentals of Probability

B.3 (i) Let Y_{it} be the binary variable equal to one if fund i outperforms the market in year t. By assumption, $P(Y_{it} = 1) = .5$ (a 50-50 chance of outperforming the market for each fund in each year). Now, for any fund, we are also assuming that performance relative to the market is independent across years. But then the probability that fund i outperforms the market in all 10 years, $P(Y_{i1} = 1, Y_{i2} = 1, \ldots, Y_{i,10} = 1)$, is just the product of the probabilities: $P(Y_{i1} = 1) \cdot P(Y_{i2} = 1) \ldots P(Y_{i,10} = 1) = (.5)^{10} = 1/1024$ (which is slightly less than .001). In fact, if we define a binary random variable Y_i such that $Y_i = 1$ if and only if fund i outperformed the market in all 10 years, then $P(Y_i = 1) = 1/1024$.

(ii) Let X denote the number of funds out of 4,170 that outperform the market in all 10 years. X has the Binomial (n,θ) distribution with $n = 4{,}170$ and $\theta = 1/1024$. The expected number of funds that will outperform the market in all 10 years is calculated as $n\theta = 4.07 \approx 4$.

(iii) Let X denote the number of funds out of 4,170 that outperform the market in all 10 years. Then $X = Y_1 + Y_2 + \ldots + Y_{4,170}$. If we assume that performance relative to the market is independent across funds, then X has the Binomial (n,θ) distribution with $n = 4{,}170$ and $\theta = 1/1024$. We want to compute $P(X \geq 1) = 1 - P(X = 0) = 1 - P(Y_1 = 0, Y_2 = 0, \ldots, Y_{4,170} = 0) = 1 - P(Y_1 = 0) \cdot P(Y_2 = 0) \cdots P(Y_{4,170} = 0) = 1 - (1023/1024)^{4170} \approx .983$. This means, if performance relative to the market is random and independent across funds, it is almost certain that at least one fund will outperform the market in all 10 years.

(iv) Using the Stata command Binomial(4170,5,1/1024), the answer is about .385. So, there is a nontrivial chance that at least five funds will outperform the market in all 10 years.

B.4 We want $P(X \geq .6)$. Because X is continuous, this is the same as $P(X > .6) = 1 - P(X \leq .6) = F(.6) = 3(.6)^2 - 2(.6)^3 = .648$. One way to interpret this is that almost 65% of all counties have an elderly employment rate of .6 or higher.

B.5 (i) As stated in the hint, if X is the number of jurors convinced of Simpson's innocence, then $X \sim \text{Binomial}(12,.20)$. We want $P(X \geq 1) = 1 - P(X = 0) = 1 - (.8)^{12} \approx .931$.

(ii) Above, we computed $P(X = 0)$ as about .069. We need $P(X = 1)$, which we obtain from (B.14) with $n = 12$, $\theta = .2$, and $x = 1$: $P(X = 1) = 12 \cdot (.2)(.8)^{11} \approx .206$. Therefore, $P(X \geq 2) \approx 1 - (.069 + .206) = .725$, so there is almost a three in four chance that the jury had at least two members convinced of Simpson's innocence prior to the trial.

B.6 E(X) = $\int_0^3 xf(x)dx$ = $\int_0^3 x[(1/9)x^2]dx$ = (1/9) $\int_0^3 x^3dx$. But $\int_0^3 x^3dx = (1/4)x^4\Big|_0^3$ = 81/4. Therefore, E(X) = (1/9)(81/4) = 9/4, or 2.25 years.

B.10 (i) E($GPA|SAT$ = 800) = .70 + .002(800) = 2.3. Similarly, E($GPA|SAT$ = 1,400) = .70 + .002(1400) = 3.5. The difference in expected GPAs is substantial, but the difference in SAT scores is also rather large.

(ii) Following the hint, we use the law of iterated expectations. Since E($GPA|SAT$) = .70 + .002 SAT, the (unconditional) expected value of GPA is .70 + .002 E(SAT) = .70 + .002(1100) = 2.9.

(iii) If a student's SAT score is 1,100, this does not mean he or she will have the GPA 2.9, found in part(ii). The GPA of 2.9 is the average GPA calculated when the average SAT score is 1,100. GPA is influenced by other variables.

. APPENDIX C
Fundamentals of Mathematical Statistics

SOLUTIONS TO PROBLEMS

C.1 (i) This is just a special case of what we covered in the text, with n = 4: E($\bar{Y}$) = μ and Var($\bar{Y}$) = $\sigma^2/4$.

(ii) E(W) = E(Y_1)/8 + E(Y_2)/8 + E(Y_3)/4 + E(Y_4)/2 = μ[(1/8) + (1/8) + (1/4) + (1/2)] = μ(1 + 1 + 2 + 4)/8 = μ, which shows that W is unbiased. Because the Y_i are independent,

$$\begin{aligned}\text{Var}(W) &= \text{Var}(Y_1)/64 + \text{Var}(Y_2)/64 + \text{Var}(Y_3)/16 + \text{Var}(Y_4)/4 \\ &= \sigma^2[(1/64) + (1/64) + (4/64) + (16/64)] = \sigma^2(22/64) = \sigma^2(11/32).\end{aligned}$$

(iii) Because 11/32 > 8/32 = 1/4, Var(W) > Var($\bar{Y}$) for any $\sigma^2 > 0$, so $\bar{Y}$ is preferred to W when each is unbiased.

C.2 (i) E(W_a) = a_1E(Y_1) + a_2E(Y_2) + … + a_nE(Y_n) = ($a_1 + a_2 + \ldots + a_n$)μ. Therefore, we must have $a_1 + a_2 + \ldots + a_n = 1$ for unbiasedness.

(ii) Var(W_a) = a_1^2 Var(Y_1) + a_2^2 Var(Y_2) + … + a_n^2 Var(Y_n) = ($a_1^2 + a_2^2 + \ldots + a_n^2$)$\sigma^2$.

(iii) From the hint, when $a_1 + a_2 + \ldots + a_n = 1$ – the condition needed for unbiasedness of W_a – we have $1/n \le a_1^2 + a_2^2 + \ldots + a_n^2$. But then $\text{Var}(\bar{Y}) = \sigma^2/n \le \sigma^2(a_1^2 + a_2^2 + \ldots + a_n^2) = \text{Var}(W_a)$.

C.3 (i) $\text{E}(W_1) = [(n-1)/n]\text{E}(\bar{Y}) = [(n-1)/n]\mu$, and so $\text{Bias}(W_1) = [(n-1)/n]\mu - \mu = -\mu/n$. Similarly, $\text{E}(W_2) = \text{E}(\bar{Y})/2 = \mu/2$, and so $\text{Bias}(W_2) = \mu/2 - \mu = -\mu/2$. The bias in W_1 tends to zero as $n \to \infty$, while the bias in W_2 is $-\mu/2$ for all n. This is an important difference.

(ii) $\text{plim}(W_1) = \text{plim}[(n-1)/n]\cdot\text{plim}(\bar{Y}) = 1\cdot\mu = \mu$ and $\text{plim}(W_2) = \text{plim}(\bar{Y})/2 = \mu/2$. Because $\text{plim}(W_1) = \mu$ and $\text{plim}(W_2) = \mu/2$, W_1 is consistent whereas W_2 is inconsistent.

(iii) $\text{Var}(W_1) = [(n-1)/n]^2\text{Var}(\bar{Y}) = [(n-1)^2/n^3]\sigma^2$ and $\text{Var}(W_2) = \text{Var}(\bar{Y})/4 = \sigma^2/(4n)$.

(iv) Because $\bar{Y}$ is unbiased, its mean squared error is simply its variance. On the other hand, $\text{MSE}(W_1) = \text{Var}(W_1) + [\text{Bias}(W_1)]^2 = [(n-1)^2/n^3]\sigma^2 + \mu^2/n^2$. When $\mu = 0$, $\text{MSE}(W_1) = \text{Var}(W_1) = [(n-1)^2/n^3]\sigma^2 < \sigma^2/n = \text{Var}(\bar{Y})$ because $(n-1)/n < 1$. Therefore, $\text{MSE}(W_1)$ is smaller than $\text{Var}(\bar{Y})$ for μ close to zero. For large n, the difference between the two estimators is trivial.

C.4 (i) Using the hint, $\text{E}(Z|X) = \text{E}(Y/X|X) = \text{E}(Y|X)/X = \theta X/X = \theta$. It follows by Property CE.4, the law of iterated expectations, that $\text{E}(Z) = \text{E}(\theta) = \theta$.

(ii) This follows from part (i) and the fact that the sample average is unbiased for the population average. Write

$$W_1 = n^{-1}\sum_{i=1}^{n}(Y_i / X_i) = n^{-1}\sum_{i=1}^{n} Z_i,$$

where $Z_i = Y_i/X_i$. From part (i), $\text{E}(Z_i) = \theta$, for all i.

(iii) In general, the average of the ratios, Y_i/X_i, is not the ratio of averages, $W_2 = \bar{Y}/\bar{X}$. (This non-equivalence is discussed a bit on page 676.) Nevertheless, W_2 is also unbiased, as a simple application of the law of iterated expectations shows. First, $\text{E}(Y_i|X_1,\ldots,X_n) = \text{E}(Y_i|X_i)$ under random sampling because the observations are independent. Therefore, $\text{E}(Y_i|X_1,\ldots,X_n) = \theta X_i$ and so

$$\text{E}(\bar{Y} \mid X_1,\ldots,X_n) = n^{-1}\sum_{i=1}^{n}\text{E}(Y_i \mid X_1,\ldots,X_n) = n^{-1}\sum_{i=1}^{n}\theta X_i$$

$$= \theta n^{-1}\sum_{i=1}^{n} X_i = \theta\bar{X}.$$

Therefore, $\text{E}(W_2 \mid X_1,\ldots,X_n) = \text{E}(\bar{Y}/\bar{X} \mid X_1,\ldots,X_n) = \theta\bar{X}/\bar{X} = \theta$, which means that W_2 is actually unbiased conditional on $(X_1,\ldots,X_n)$ and therefore also unconditionally unbiased.

(iv) For the n = 17 observations given in the table – which are, incidentally, the first 17 observations in the file CORN.RAW – the point estimates are w_1 = .418 and w_2 = 129.09/308.76 = .418. The estimates are same for this example. If we use w_1, we estimate $E(Y|X = x)$ for any $x > 0$ as $E(Y \mid X = x)$ = .418 x. For example, if x = 300, then the predicted yield is .418(300) = 125.4.

C.5 (i) While the expected value of the numerator of G is $E(\bar{Y}) = \theta$, and the expected value of the denominator is $E(1 - \bar{Y}) = 1 - \theta$, the expected value of the ratio is not the ratio of the expected value.

(ii) By Property PLIM.2(iii), the plim of the ratio is the ratio of the plims (provided the plim of the denominator is not zero): $\text{plim}(G) = \text{plim}[\bar{Y}/(1 - \bar{Y})] = \text{plim}(\bar{Y})/[1 - \text{plim}(\bar{Y})] = \theta/(1 - \theta) = \gamma$.