ISQS 5347 Midterm 2
Closed book and notes, no electronic devices. Points (out of 100) in parentheses.

1.(20) The time (T) it takes to serve a customer in line has mean 0.5 minutes and standard deviation 1.0 minutes. Find the mean, variance, and standard deviation of the total time ($T_1+T_2+\ldots+T_{100}$) that it takes to serve the next 100 customers. Assume independence and provide all justifications such as linearity, additivity, algebra, etc.

Solution: The mean of the total time calculated as

$E(T_1+T_2+\ldots+T_{100})$
$= E(T_1)+E(T_2)+\ldots+ E(T_{100})$ (by the additivity property of expectation)
$= 0.5 + 0.5 +\ldots+ 0.5$ (assuming identical distributions, as suggested in the problem statement)
$= 100(0.5)$ (because there are 100 summands)
$= 50.$

The variance of the total time is calculated as

$Var(T_1+T_2+\ldots+T_{100})$
$= Var(T_1)+Var(T_2)+\ldots+ Var(T_{100})$ (by the additivity property of variance under independence)
$= 1.0 + 1.0 +\ldots+ 1.0$ (assuming identical distributions, as suggested in the problem statement)
$= 100(1.0)$ (because there are 100 summands)
$=100.$

The standard deviation of the total time is, by definition, the square root of the variance of the total time, and is therefore equal to 10.

2.A.(15) Find the mean, variance, and standard deviation of a discrete random variable that has the following distribution.

y	$p(y)$
0	0.8
10	0.2

Solution: Using the formula for the mean of a discrete random variable, the mean of the random variable Y is given by $\mu = \Sigma y\, p(y) = (0)(0.8) + (10)(0.2) = 2.0$.

Using the formula for the variance of a discrete random variable, the variance of the random variable Y is given by $\sigma^2 = \Sigma(y-\mu)^2\, p(y) = (0-2)^2(0.8) + (10-2)^2(0.2) = 3.2 + 12.8 = 16.0$.

Because the standard deviation is the square root of the variance, the standard deviation of the random variable Y is given by $\sigma = \sqrt{16.0} = 4.0$.

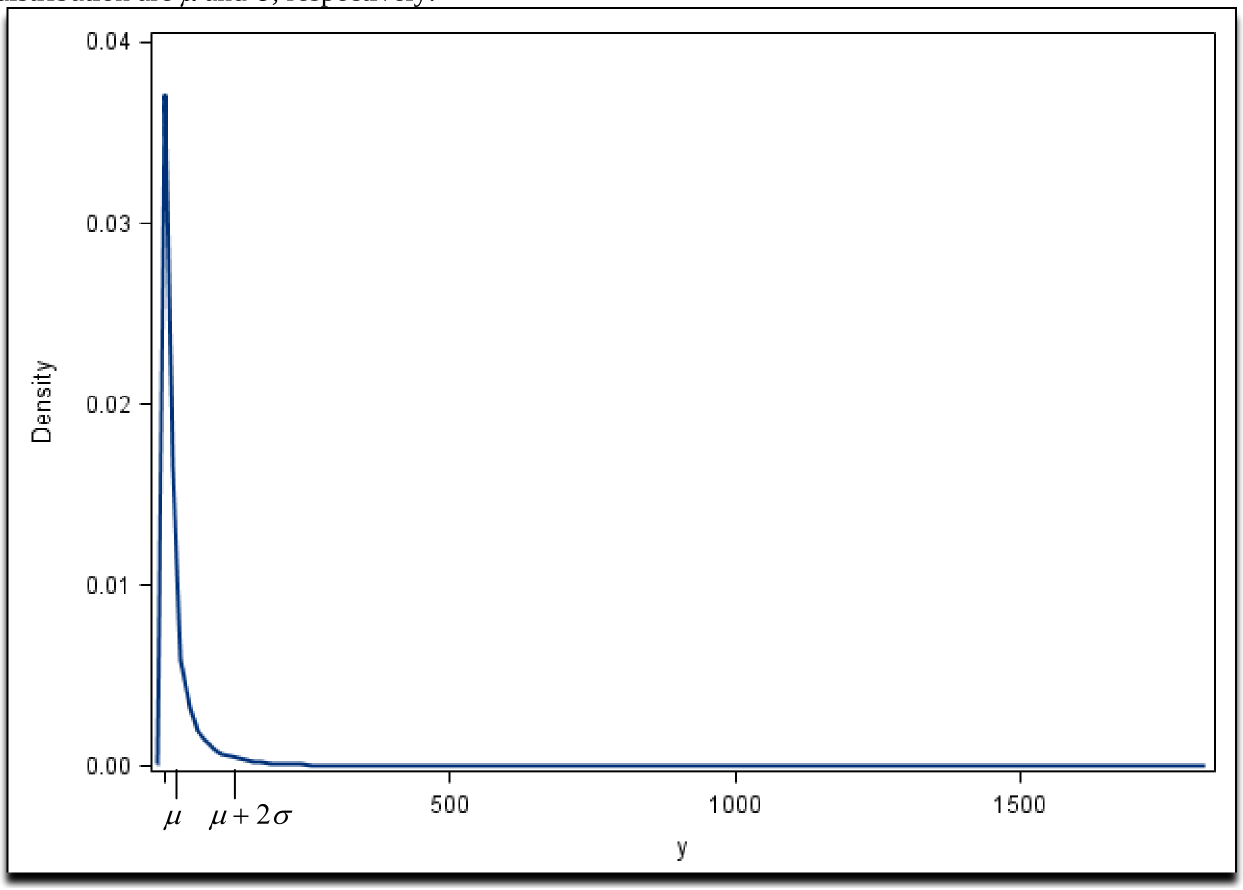
2.B.(7) Suppose the distribution of 2.A. produces an iid sample Y_1, Y_2, …, $Y_{10{,}000}$. Explain what the Law of Large Numbers tells you about $\bar{Y} = \frac{1}{10{,}000}\sum_{i=1}^{10{,}000} Y_i$. Be specific.

Solution: The Law of Large Numbers (LLN) states that the average of an iid sample that is produced by a distribution $p(y)$ approaches the expected value (either $\Sigma y\, p(y)$ or $\int y\, p(y)dy$, but $\Sigma y\, p(y)$ in this case) as the sample size gets larger. Here the sample size is pretty large, at n=10,000. So, by the LLN, the sample average $\bar{Y}$ will be close to 2.0.

2.C.(13) Suppose the distribution of 2.A. produces an iid sample Y_1, Y_2, …, $Y_{10{,}000}$. Explain what the Central Limit Theorem tells you about $\bar{Y} = \frac{1}{10{,}000}\sum_{i=1}^{10{,}000} Y_i$. Be specific.

Solution: The Central Limit Theorem states that the distribution of $\bar{Y}$ will be approximately a normal distribution when the sample size is large. Here, the mean of the distribution of $\bar{Y}$ is 2.0 and the standard deviation is $4.0/\sqrt{10{,}000} = 4.0/100 = 0.04$. Hence, by the CLT, $\bar{Y}$ will be within the range 1.92 and 2.08 with approximately 95% probability.

3. The distribution that produces the data is shown below. The mean and standard deviation of the distribution are μ and σ, respectively.



3.A.(10) What does Chebychev's inequality tell you about the percentage of the observations (Y) that are *larger than* $\mu + 2\sigma$?

Solution: As discussed in class, Chebychev's inequality states that *at least* 75% of the observations are within the range ($\mu - 2\sigma$, $\mu + 2\sigma$). Hence, *at most* 25% of the observations can be outside the range. Chebychev's inequality does not tell us how the numbers outside the ($\mu - 2\sigma$, $\mu + 2\sigma$) range are distributed into in the upper and lower tail; hence, all we can say is that no more than 25% of the observations can be larger than $\mu + 2\sigma$.

3.B.(20) The distribution produces iid data $Y_1, Y_2, \ldots, Y_n$, and

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad , \qquad \hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

are used to estimate μ and σ. Is $\bar{Y} + 2\hat{\sigma}$ an unbiased estimator of $\mu + 2\sigma$? Use linearity, additivity and Jensen's inequality *separately* in your answer. (In other words, don't combine the three concepts in your explanation. Use each concept separately at distinct places.)

Solution: By definition, an estimator is unbiased if the expected value of the estimator is equal to the estimand. Hence, the estimator $\bar{Y}+2\hat{\sigma}$ an unbiased estimator of $\mu + 2\sigma$ if $E(\bar{Y}+2\hat{\sigma})=\mu+2\sigma$. To check whether this is true, consider

$E(\bar{Y}+2\hat{\sigma})=E(\bar{Y})+E(2\hat{\sigma})$ (by the additivity property of expectation)

$=E(\bar{Y})+2E(\hat{\sigma})$ (by the linearity property of expectation)

$=\mu+2E(\hat{\sigma})$ (since $\bar{Y}$ is an unbiased estimator of μ under the assumptions given).

Thus, if $\hat{\sigma}$ were an unbiased estimator of σ, it would then follow that $\bar{Y}+2\hat{\sigma}$ is an unbiased estimator of $\mu + 2\sigma$.

But $\hat{\sigma}$ is not an unbiased estimator of σ. We know from class discussions that $\hat{\sigma}^2=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2$ is an unbiased estimator for σ^2, implying that $E(\hat{\sigma}^2)=\sigma^2$. However, $\hat{\sigma}=(\hat{\sigma}^2)^{1/2}$ is a nonlinear function, so the expected value of the function is not equal to the function of the expected value. Here, the function is the square root function $f(x)=x^{1/2}$, which is concave since $f'(x)=(1/2)\,x^{-1/2}$, implying $f''(x)=(-1/4)\,x^{-3/2}$, which is less than zero for all $x>0$. Hence, by Jensen's inequality, $E(\hat{\sigma})=E\{(\hat{\sigma}^2)^{1/2}\}<\{E(\hat{\sigma}^2)\}^{1/2}=\{\sigma^2\}^{1/2}=\sigma$, implying that, in turn, $E(\bar{Y}+2\hat{\sigma})<\mu+2\sigma$, and that $\bar{Y}+2\hat{\sigma}$ is a biased estimator of $\mu + 2\sigma$.

4.(15) The data $y_1 = 1/2$ and $y_2 = 2/3$ are obtained as an iid sample from $p(y)$, where

$$p(y)=\theta y^{\theta-1}\ , \text{ for } 0< y < 1.$$

Find the MLE for θ. Show your work.

Solution: The likelihood function for an iid sample is the product of the pdfs:

$L(\theta\,|\,y_1, y_2) = \theta(1/2)^{\theta-1}\times \theta(2/3)^{\theta-1} = \theta^2(1/2 \times 2/3)^{\theta-1} = \theta^2(1/3)^{\theta-1}$.

The maximum likelihood estimate is the value $\hat{\theta}$ that maximizes $L(\theta\,|\,y_1, y_2)$. It is easier to find the $\hat{\theta}$ that maximizes $\ln(L(\theta\,|\,y_1, y_2))$, and the value $\hat{\theta}$ that maximizes $\ln(L(\theta\,|\,y_1, y_2))$ is the same $\hat{\theta}$ that maximizes $L(\theta\,|\,y_1, y_2)$, because the $\ln(x)$ function is monotonically increasing.

So we consider $\text{LL}(\theta) = \ln(L(\theta\,|\,y_1, y_2)) = \ln(\theta^2(1/3)^{\theta-1}) = 2\ln(\theta) + (\theta-1)\ln(1/3) = 2\ln(\theta) - (\theta-1)\ln(3)$, by properties of logarithms.

Finding the derivative, and solving for the MLE $\hat{\theta}$ that makes the derivative zero, we have

$\text{LL}'(\hat{\theta}) = 0$, or

$2/\hat{\theta} - \ln(3) = 0$, or

$2/\hat{\theta} = \ln(3)$, or

$\hat{\theta} = 2/\ln(3)$.