Closed book and notes. 120 minutes.

1. True or false. (for each, 2 points if correct, 1 point if left blank.)

(a) T F When choosing a random sample without replacement from a finite population, it is possible that some member of the population is chosen twice.

(b) T F For any random variables X_1 and X_2, the following result is true: $E(X_1+X_2) = E(X_1) + E(X_2)$.

(c) T F A point estimator is often a single number, but it can be an interval.

(d) T F A point estimator is said to be "maximum likelihood" if it is equal to the largest value in the sample.

(e) T F Chebyshev's Inequality guarantees that the sum of many random variables is (at least approximately) normally distributed.

(f) T F The standard deviation of a point estimator is called its standard error.

(g) T F If $\hat{\Theta}$ is an unbiased estimator, then the mean squared error of $\hat{\Theta}$ is equal to the variance of $\hat{\Theta}$.

(h) T F Mean squared error is one, but not the only, measure of a confidence interval's quality.

(i) T F Let $\bar{X}$ denote the sample mean of a random sample of size n. The standard error of $\bar{X}$ decreases as the sample size n increases.

(j) T F The ith observation from a sample of size n is called the ith order statistic.

(k) T F A statistic is a function of the observed sample values.

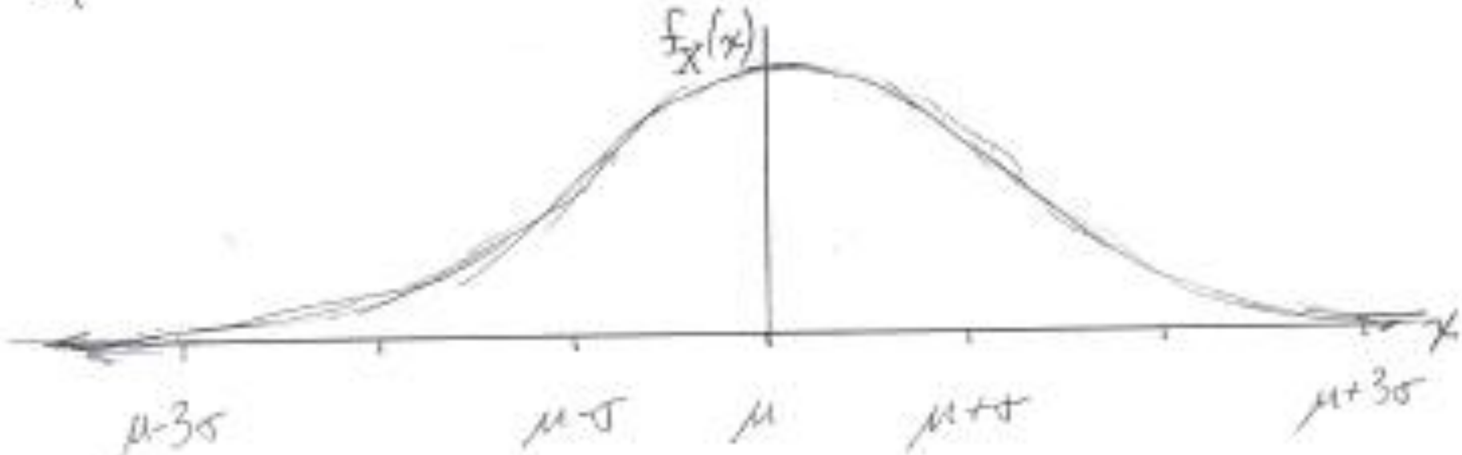
(l) T F If (X, Y) has a bivariate normal distribution, then $P(X < 0, Y < 0) = 0$.

(m) T F If X and Y are independent, then $\text{corr}(X, Y) = 0$.

(n) T F Let X have a binomial distribution with $n = 5$ and $p = .2$. Without the continuity correction, the normal approximation to the binomial would yield $P(X = 3) = 0$.

2. Let X denote a normally distributed random variable with mean μ and standard deviation σ. The pth quantile of X is the constant x_p that satisfies $P(X \le x_p) = p$. When $\mu = 0$ and $\sigma = 1$, the corresponding value is denoted by z_p.

6 pt. (a) Sketch the density function of X. Label both axes. Scale the horizontal axis.



4 pt (b) Circle the largest value.

$x_{.01}$ x_0 0 $x_{.99}$

(Answer depends on the values of μ & σ, so either "0" or "$x_{.99}$" is ok.)

4 pt. (c) Circle the true statement.

$z_p = (x_p - \mu)/\sigma$ $z_p = \mu + \sigma x_p$ $z_p = x_p$ $z_p = x_p^{0.135}$

3. Throughout this course we have followed a consistent notational convention that indicates the nature of various quantities. Describe each expression below by writing "random variable", "event", "constant", or "undefined" on the blank lines.

2 pts each

(a) X^2 RV

(b) $Y < X$ Event

(c) $\mathrm{Var}(Y < X)$ Undefined

(d) $E(X)$ constant

(e) 0 constant

(f) σ constant

(g) $P(X = 3)$ constant

4. Suppose that you invest \$4000 in IBM (ticker: IBM), a large hardware and software company, and \$6000 in Symix Systems (ticker: SYMX), a small software company. Let $V = \$4000X + \$6000Y$ be the portfolio value one year from now. That is, let X denote the relative change in IBM stock price and Y the relative price change for SYMX. Assume that $\mu_X = 1.2$, $\mu_Y = 1.3$, $\sigma_X = 0.2$, $\sigma_Y = 0.4$, and corr$(X, Y) = 0.6$.

4 pt (a) Are X and Y independent? Yes No $(\text{corr}(X,Y) \neq 0)$

6 pt (b) Find the covariance of X and Y. (Recall that the correlation is the covariance divided by the standard deviations.)

$$\text{Cov}(X,Y) = \text{Corr}(X,Y)\,\sigma_X\,\sigma_Y$$
$$= (.6)(.2)(.4)$$
$$= .048 \longleftarrow$$

6 pt (c) Find E(V). (Recall that the expected value of a sum is the sum of expected values.)

$$E(V) = E(\$4000X + \$6000Y)$$
$$= \$4000\,E(X) + \$6000\,E(Y)$$
$$= \$4000(1.2) + \$6000(1.3)$$
$$= \$12{,}600 \longleftarrow$$

6 pt (d) Find Var(V). (Recall that the variance of a sum is the sum of the covariances.)

$$\text{Var}(V) = \text{Var}(\$4000X + \$6000Y)$$
$$= \$4000^2\,\text{Var}(X) + 2(\$4000)(\$6000)\,\text{Cov}(X,Y) + \$6000^2\,\text{Var}(Y)$$
$$= \$4000^2(.2)^2 + 2(\$4000)(\$6000)(.048) + \$6000^2(.4)^2$$
$$= (.64 + 2.304 + 5.76) \times (10^6\,\$^2)$$
$$= (\$^2)\,8{,}704{,}000 \longleftarrow$$

5. We want to estimate p, the probability of success for a sequence of Bernoulli trials. We repeatedly (and independently) observe the geometric random variable X, the number of trials to obtain a success; let $X_1, X_2, \ldots, X_n$ denote n such observations.

(If you wish to consider a specific example, consider the "penny-tipping" exercise, with each student tipping a penny until it lands heads up. Let X_i be the number of tips required by student i, for $i = 1, 2, \ldots, 130$. We want to estimate p, the probability that the penny lands heads up on any one trial.)

5 pt (a) What are the possible values of X_1?

$$\{1, 2, \ldots\}$$

6 pt (b) Write the mass function of X_1 in terms of p.

X_i is geometric and identically distributed

$$f_X(x) = \begin{cases} p(1-p)^{x-1} & \text{if } x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

6 pt (c) Write the likelihood function of the sample of size n for $p = .5$.

$$L(p) = \prod_{i=1}^{n} f_X(x_i; p)\Big|_{p=.5}$$
$$= \prod_{i=1}^{n} (.5)(1-.5)^{x_i - 1}$$
$$= (.5)^{\sum_{i=1}^{n} x_i}$$

any are ok.

5 pt (d) Exactly one of the following statistics is the maximum-likelihood estimator of p. Circle that estimator. (You do not need to derive the MLE. Rather, simply eliminate the choices that are not reasonable. Here $\bar{X}$ denotes the sample average.)

$\hat{p} = \bar{X}$ $\qquad$ $\hat{p} = \dfrac{1}{\bar{X}}$ $\qquad$ $\hat{p} = \prod_{i=1}^{n} X_i$

 1st Name Key

6. (Montgomery and Runger, 6-32.) In the transmission of digital information, the probability that a bit has high, moderate, or low distortion is 0.01, 0.04 and 0.95, respectively. Suppose that three bits are transmitted and that the amount of distortion of each bit is independent of the other bits.

p_1 p_2 p_3

5 pt (a) Consider a fourth outcome: that a bit has *no* distortion. What is the probability that the first bit has no distortion?

$$P(\text{no distortion}) \le 1 - 0.01 - 0.04 - 0.95$$
$$\le 0$$
$$\Rightarrow P(\text{no distortion}) = 0$$

6 pt (b) What is the probability that exactly two bits have high distortion and one has moderate distortion?

$$P(X_1 = 2, X_2 = 1, X_3 = 0) = \binom{3}{2,1,0}(.01)^2(.04)^1(.95)^0$$
$$= 3(.01)^2(.04)$$
$$= .000012$$

because (X_1, X_2, X_3) is multinomial.

6 pt (c) What is the expected number of bits having low distortion?

$$E(X_3) = n p_3$$
$$= 3(.95)$$
$$= 2.85$$

6 pt (d) Conditional that the first bit has low distortion, what is the probability that the second bit has high distortion?

0.95

because bits are independent

7. (Montgomery and Runger, 3-106.) Suppose that a lot of washers is large enough that it can be assumed that the sampling is done with replacement. Assume that 60% of the washers exceed a target thickness. Let A_i denote the event that washer i exceeds the target thickness.

Given: $P(A_i) = .6$ for $i = 1, 2, \ldots$

5 pt. (a) What is the probability that a randomly selected washer does not exceed its target thickness?

$$P(A_i') = 1 - P(A_i)$$
$$= 1 - .6$$
$$= .4$$

5 pt. (b) Write the event (not its probability) that washers 1, 2, and 3 all exceed the target thickness.

$$A_1 \cap A_2 \cap A_3$$

7 pt. (c) What is the minimum number of washers, n, that need to be selected so that the probability that all the washers are thinner than the target is less than 0.10?

$$P(A_1' \cap A_2' \cap \cdots \cap A_n') < 0.10$$
$$\Rightarrow P(A_1') P(A_2') \cdots P(A_n') < 0.10 \quad \text{(independent)}$$
$$\Rightarrow (.4)^n < 0.10 \quad \text{(Part a)}$$
$$\Rightarrow n < (\ln .10)/(\ln .4)$$
$$\Rightarrow n = 3$$

5 pt. (d) Circle the correct answer. The first sentence of this problem (where "with replacement" is assumed) affects the answer to Part(s)

(a) (b) (c) (a,b) (a,c) (b,c) (a,b,c)

"Replacement" $\Rightarrow$ independence

8. Let X denote the result of rolling a six-sided die; that is, X is the number of dots facing up. Assume that all six sides are equally likely.

6 pt (a) Write the mass function of X? (Be complete.)

$$f_X(x) = \begin{cases} 1/6 & \text{if } x = 1,2,3,4,5,6 \\ 0 & \text{otherwise} \end{cases}$$

6 pt (b) Find $E(X^2)$.

$$E(X^2) = 1^2\left(\tfrac{1}{6}\right) + 2^2\left(\tfrac{1}{6}\right) + \cdots + 6^2\left(\tfrac{1}{6}\right)$$
$$= 15\tfrac{1}{6} \leftarrow$$

6 pt (c) Find the conditional mass function of X given that $X < 3$. (Be complete.)

$$X < 3 \Rightarrow X \in \{1, 2\} \Rightarrow \boxed{P(X = x) = 0 \text{ elsewhere}}$$

For $x = 1, 2$

$$P(X = x \mid X < 3) = \frac{P(X = x \text{ and } X < 3)}{P(X < 3)}$$
$$= \frac{P(X = x)}{P(X < 3)}$$
$$= \frac{1/6}{1/3} = \frac{1}{2} \leftarrow \boxed{\text{for } x = 1, 2}$$

6 pt (d) Find $E(X \mid X < 2)$.

$$X < 2 \Rightarrow P(X = 1 \mid X < 2) = 1$$
$$\Rightarrow E(X \mid X < 2) = 1 \leftarrow$$

9. (Montgomery and Runger, 5-72.) Suppose that the log-ons to a computer network follow a Poisson process with an average of three log-ons per minute.

6 pt (a) What is the mean time between log-ons?

$E(\text{time between log-ons}) = 1/\lambda = 1/3$ minute ←

6 pt (b) What is the probability that exactly two log-ons occur during a particular two-minute time interval?

X = "# log-ins in two minutes"

$\Rightarrow X \sim$ Poisson (mean = 6 log-ons)

$\Rightarrow P(X=2) = \dfrac{e^{-6}\, 6^2}{2!}$

$\approx .0446$ ←

6 pt (c) Determine the time t (in minutes) so that the probability that at least one log-on occurs before t minutes is 0.95.

Let T = "time to next log-on"

Then $T \sim$ exponential (mean = 1/3 min)

$\Rightarrow P(T < t) = .95$

$\Rightarrow 1 - e^{-3t} = .95$ (cdf of exponential)

$\Rightarrow t = \dfrac{-\ln(.05)}{3} \approx .9986$ ←

5 pt (d) Let Y have an Erlang distribution with shape parameter $r = 3$ and $\lambda = 3$. In the context of this problem, what might be the physical meaning of Y?

Let Y = "time of the third log-in"