ISyE 6644 — Spring 2019 — Test #3 Solutions

This test is 120 minutes. You're allowed *three* cheat sheets (6 sides total).

This test requires a proctor. All questions are 3 points, except #31, which is *10 points*.

1. TRUE or FALSE? If $f(x, y) = c/(x + y)$ for all $0 < x < 1$ and $1 < y < 2$, where c is whatever value makes this thing integrate to 1, then X and Y are independent random variables.

   Solution: FALSE. (Because $f(x, y)$ doesn't factor properly.) $\square$

2. TRUE or FALSE? In Arena, the `Queue` spreadsheet in the Advanced Process panel can be used to change a queue's discipline from FIFO to LIFO.

   Solution: FALSE. (It's in the Basic Process panel.) $\square$

3. What does an Arena `Dispose` module look like?

   (a) Rectangle
   (b) Oval
   (c) Right-pointing trapezoid
   (d) Left-pointing trapezoid
   (e) Diamond

   Solution: (d). $\square$

4. Suppose an arrival occurs in a discrete-event simulation. What does the simulation typically do at that point?

   (a) It skips ahead to the next arrival time.
   (b) It schedules the next arrival event on the future events list.
   (c) It turns on the next server.

(d) It schedules the arriving customer's departure time.

(e) It removes one customer from the FIFO queue.

Solution: (b). $\square$

5. TRUE or FALSE? The Kolmogorov-Smirnov test can be used to see if data seem to fit to a particular hypothesized distribution.

   Solution: TRUE. (It's a goodness-of-fit test.) $\square$

6. TRUE or FALSE? If X_1 and X_2 are i.i.d. Nor(0,4) random variables, then the ratio X_1/X_2 has both the Cauchy distribution and the t distribution with one degree of freedom.

   Solution: TRUE.

$$\frac{\text{Nor}(0,4)}{\text{Nor}(0,4)} \sim \frac{2\,\text{Nor}(0,1)}{2\,\text{Nor}(0,1)} \sim \text{Cauchy}.$$

   And we also mentioned in class that the Cauchy is a special case of the t with one d.f. $\square$

7. Suppose $X_1, X_2, \ldots$ is a stationary process with mean μ and variance parameter $\sigma^2 \equiv \lim_{n\to\infty} n\mathsf{Var}(\bar{X})$, where the sample mean $\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i$. What colorful stochastic process does $\sum_{i=1}^{\lfloor nt \rfloor}(X_i - \mu)/(\sigma\sqrt{n})$ converge to (as a function of t)?

   (a) Redian motion

   (b) Yellowian motion

   (c) Brownian motion

   (d) Greenian motion

   Solution: (c). $\square$

8. Find the sample variance of 0, 10, 20.

(a) 0
(b) 10
(c) 100
(d) 200
(e) None of the above

Solution: $S^2 = 100$. So (c) is the answer. $\square$

9. If $X_1, \ldots, X_n$ are i.i.d. Geom(1/3), what is the expected value of the sample mean $\bar{X}$?

(a) 1
(b) 3
(c) 9
(d) 1/3
(e) 1/9

Solution: Since the sample mean is always unbiased for the true mean, we have $\mathsf{E}[\bar{X}] = \mathsf{E}[X_i] = 1/p = 3$, so the answer is (b). $\square$

10. If $X_1, \ldots, X_n$ are i.i.d. Bin(3, 0.6), what is the expected value of the sample variance S^2?

(a) 0.72
(b) 1.2
(c) 1.44
(d) 1.8
(e) None of the above

Solution: Since the sample variance is always unbiased for the true variance, we have $\mathsf{E}[S^2] = \mathsf{Var}(X_i) = npq = 0.72$, so the answer is (a). $\square$

11. TRUE or FALSE? If T_1 is an unbiased estimator for some parameter θ and T_2 is biased as an estimator for θ, then T_1 may or may not have smaller MSE than T_2.

Solution: TRUE. $\mathsf{MSE} = \mathsf{Bias}^2 + \mathsf{Var}$, so the MSE competition could go either way depending on the magnitudes of the biases and variances. $\square$

12. If $X_1, \ldots, X_5$ are i.i.d. Nor$(5, 20)$, what is the expected value of the maximum likelihood estimator for the variance σ^2?

    (a) 5
    (b) 16
    (c) 20
    (d) 25
    (e) 100
    (f) None of the above

Solution: We know from class that for i.i.d. normal observations, the MLE $\widehat{\sigma^2} = \frac{n-1}{n} S^2$; and we also remember that S^2 is unbiased for σ^2. Thus,

$$\mathsf{E}[\widehat{\sigma^2}] \;=\; \frac{n-1}{n}\,\mathsf{E}[S^2] \;=\; \frac{n-1}{n}\,\sigma^2 \;=\; \frac{4}{5} \times 20 \;=\; 16.$$

So the answer is (b). $\square$

13. Suppose we observe the Bern(p) realizations $X_1 = 0$, $X_2 = 0$, $X_3 = 1$, $X_4 = 1$, and $X_5 = 0$. What is the maximum likelihood estimate of p?

    (a) 0
    (b) 0.24
    (c) 0.4
    (d) 0.6
    (e) 1
    (f) None of the above

Solution: $\hat{p} = \bar{X} = 2/5$. So (c). $\square$

14. Suppose $X_1, \ldots, X_6$ are i.i.d. Bern(p), and we obtain the maximum likelihood estimate $\hat{p} = 0.5$. What's the maximum likelihood estimate of $\mathsf{Var}(X_i) = p(1-p)$?

    (a) 0
    (b) 0.125
    (c) 0.25
    (d) 0.5
    (e) 1
    (f) Not enough data to tell

    Solution: By invariance,

    $$\widehat{\mathsf{Var}}(X_i) \;=\; \hat{p}(1-\hat{p}) \;=\; 0.25.$$

    Thus, (c) is correct. $\square$

15. Suppose we're conducting a χ^2 goodness-of-fit test to determine whether or not 1000 i.i.d. observations are from a shifted Gamma(α, β, c), where α, β, and the shift parameter c must all be estimated. If we divide the observations into 15 equal-probability intervals, how many degrees of freedom will our test have?

    (a) 3
    (b) 11
    (c) 14
    (d) 15
    (e) 996
    (f) 999

    Solution: Let $s = 3$ denote the number of parameters that must be estimated using $k = 15$ intervals. Then $\nu = k - s - 1 = 15 - 3 - 1 = 11$. This is answer (b). $\square$

16. TRUE or FALSE? Newton's method can help you find the zeros of a continuous function $g(x)$, but you need knowledge of the derivative $g'(x)$.

    Solution: TRUE. $\square$

17. Generally speaking, we can break simulation *output analysis* problems into two categories: Steady-state simulations and __________ simulations. Fill in the blank.

    (a) duscrete-event

    (b) continuous-time

    (c) terminating

    (d) truncated

    (e) nominal

    Solution: (c). □

18. Consider the following 8 observations arising from a simulation:

    154 180 175 162 200 173 191 183

    Use the method of batch means to calculate a two-sided 90% confidence interval for the mean μ. In particular, use two batches of size four.

    (a) [54, 300]

    (b) [117, 237]

    (c) [154, 200]

    (d) [170, 184]

    (e) 177 ± 30

    Solution: The batch size is $m = 4$, the number of batches is $b = 2$, and the total number of observations is $n = 8$. The grand sample mean is $\bar{X} = 177.25$. The batch means are

    $$\bar{X}_{1,4} = 167.75 \quad \text{and} \quad \bar{X}_{2,4} = 186.75.$$

    The batch means variance estimator is

    $$\widehat{V}_B = \frac{m}{b-1}\sum_{i=1}^{b}(\bar{X}_{i,m} - \bar{X})^2 = 722.$$

The batch means confidence interval is

$$\begin{aligned}\mu &\in \bar{X} \pm t_{\alpha/2,b-1}\sqrt{\widehat{V}_B/n} \\ &= 177.25 \pm t_{0.05,1}\sqrt{722/8} \\ &= 177.25 \pm 59.98 \;=\; [117.27,\, 237.23].\end{aligned}$$

So the answer is (b). $\square$

19. Consider a particular data set of 30000 stationary waiting times obtained from a large queuing system. Suppose your goal is to get a confidence interval for the unknown mean. Would you rather use (a) 30 batches of 1000 observations or (b) 3000 batches of 10 observations each?

    Solution: (a), since that choice has proper asymptotic properties. $\square$

20. Suppose $[0,4]$ is a 95% nonoverlapping batch means confidence interval for the mean μ based on 20 batches of size 500. Now the boss has decided that she wants a 90% CI based on those same 20 batches of size 500. What is it?

    (a) [0, 4]
    (b) [0.35, 3.65]
    (c) [1, 3]
    (d) [1.35, 2.65]
    (e) 2 ± 3

    Solution: The confidence interval is of the form

$$[0,4] \;=\; \bar{X} \pm t_{\alpha/2,b-1}\sqrt{\widehat{V}_B/n}.$$

This implies that $\bar{X} = 2$ and the half-length is $t_{0.025,19}\sqrt{\widehat{V}_B/n} = 2$. Thus, the new 90% confidence interval is

$$\begin{aligned}\text{new CI} &= \bar{X} \pm t_{0.05,19}\sqrt{\widehat{V}_B/n} \\ &= 2 \pm \frac{t_{0.05,19}}{t_{0.025,19}} t_{0.025,19}\sqrt{\widehat{V}_B/n} \\ &= 2 \pm \frac{1.729}{2.093} \times 2 \\ &= 2 \pm 1.652 \;=\; [0.348,\, 3.652].\end{aligned}$$

Thus, (b) is correct. $\square$

21. Consider the following 5 observations:

$$54 \quad 80 \quad 75 \quad 62 \quad 90$$

If we choose a batch size of 4, calculate all of the overlapping batch means for me.

(a) 72.2
(b) 75
(c) 67.75, 76.75
(d) 69.7, 72.3, 75.7

Solution: $\bar{X}^{\text{o}}_{1,4} = \frac{1}{4}\sum_{i=1}^{4} X_i = 67.75$ and $\bar{X}^{\text{o}}_{2,4} = \frac{1}{4}\sum_{i=2}^{5} X_i = 76.75$. Therefore, the answer is (c). $\square$

22. TRUE or FALSE? Suppose that $X_1, X_2, \ldots$ is a stationary stochastic process with covariance function $R_k \equiv \mathsf{Cov}(X_1, X_{1+k})$, for $k = 0, 1, \ldots$. Then the variance of the sample mean can be represented as

$$\mathsf{Var}(\bar{X}) = \frac{1}{n}\left[R_0 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)R_k\right].$$

Solution: TRUE. $\square$

23. TRUE or FALSE? Using the notation of the previous question, $\lim_{n\to\infty} n\mathsf{Var}(\bar{X}) = \sum_{i=0}^{\infty} R_i$.

Solution: FALSE — $\lim_{n\to\infty} n\mathsf{Var}(\bar{X}) = \sum_{i=-\infty}^{\infty} R_i = R_0 + 2\sum_{i=1}^{\infty} R_i$. $\square$

24. Consider the output analysis method of nonoverlapping batch means. Assuming that you have a sufficiently large batch size, it can be shown (take my honest word for it!) that when the number of batches b is *even*, the *expected width* of the 90% two-sided confidence interval for μ is proportional to

$$\frac{t_{0.05,b-1}}{\sqrt{b-1}} \frac{\left(\frac{b-1}{2}\right)\left(\frac{b-3}{2}\right)\cdots\frac{1}{2}}{\left(\frac{b-2}{2}\right)!}.$$

Using the above equation, determine which of (i) $b = 2$ or (ii) $b = 6$ gives the smaller expected width.

Solution: Let $h(b)$ denote the value of the above expression as a function of b. Then easy calculations reveal that $h(b) = 3.157$ and $h(6) = 0.845$. So the answer is (ii). □

25. Which variance reduction method is most-closely associated with a paired-t confidence interval for the mean?

    (a) common random numbers
    (b) antithetic random numbers
    (c) control variates
    (d) composition

    Solution: (a). □

26. Which variance reduction method takes the average of two negatively correlated estimators for the mean to get a lower-variance estimator for the mean?

    (a) common random numbers
    (b) antithetic random numbers
    (c) control variates
    (d) composition

    Solution: (b). □

27. What are ranking and selection methods most well known for?

    (a) Finding the best of a number of competing systems
    (b) Finding a confidence interval for the mean
    (c) Finding a CI for the variance
    (d) Estimating the power of a hypothesis test

(e) Determining a good truncation point for a steady-state simulation.

Solution: (a). $\square$

28. Suppose we are interested in determining which of 5 potential inventory policies has the highest probability of making a profit for us this year. Which type of ranking and selection problem is this?

    (a) Normal

    (b) Bernoulli

    (c) Poisson

    (d) Exponential

    Solution: (b). $\square$

29. Consider a normal ranking and selection problem in which we are trying to determine which of three (simulated) strategies maximizes our expected profit. After the R&S procedure finishes, we have the following sample profits: $\bar{X}_1 = 100$, $\bar{X}_2 = 500$, and $\bar{X}_3 = -100$. Which system do you choose as best?

    (a) System 1

    (b) System 2

    (c) System 3

    Solution: System 2 — the one with the largest sample mean (duh)! So (b). $\square$

30. Let's taste test to determine which of Coke vs. Pepsi is the more-preferred by Atlantans. Without going into the details regarding the parameter choices for $P^\star$ and $\delta^\star$, let's just suppose that the single-stage multinomial ranking-and-selection procedure from class tells us to survey 1500 people. But after just 1000 people, suppose that 751 love Coke, while only 249 love Pepsi. What do you do?

    (a) You are stubborn and inefficient — you take all 1500 samples even though Pepsi cannot possibly catch up.

(b) You are smart and efficient — since the R&S procedure will select the soft drink based solely on which one gets more wins, you stop now, select Coke as the winner, and save 500 expensive observations!

(c) You go to UGA and you burp a lot.

Solution: (b). □

31. (10 points) Who is the best, nicest, most-adorable teacher you've ever had?

   (a) Dave Goldsman

   Solution: (a). □