NAME →

ISyE 3044 — Test 3 Solutions — Spring 2013

(revised 12/5/13)

This test is 105 minutes. You're allowed three cheat sheets. **Just show your extremely neat answers — any intermediate steps, multiple answers, or untidiness will be penalized.** All questions are 3 points, except #1 (1 fabulous point). Good luck!

1. What is the secret word?

   Solution: Loyalty. $\diamondsuit$

2. Use Monte Carlo simulation with PRN's $U_1 = 0.75$, $U_2 = 0.21$, $U_3 = 0.57$, and $U_4 = 0.39$ to approximate $\int_0^3 e^{-x^2/9}\,dx$.

   Solution:

$$\hat{I}_n \;=\; \frac{b-a}{n}\sum_{i=1}^{n} f(a+(b-a)U_i) \;=\; \frac{3}{4}\sum_{i=1}^{4} f(3U_i) \;=\; \frac{3}{4}\sum_{i=1}^{4} e^{-U_i^2} \;=\; 2.332. \quad \diamondsuit$$

3. If X and Y have joint p.d.f. $f(x,y) = xy$, $0 \le x \le y \le c$, for some appropriate constant c, find $\mathsf{E}[X]$.

   Solution: First, get c.

$$1 \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy \;=\; \int_0^c \int_0^y xy\,dx\,dy \;=\; \frac{c^4}{8},$$

   so that $c = 8^{1/4}$. Now,

$$f_X(x) \;=\; \int_{-\infty}^{\infty} f(x,y)\,dy \;=\; \int_x^{8^{1/4}} xy\,dy \;=\; \frac{\sqrt{8}\,x - x^3}{2}, \quad 0 < x < 8^{1/4}.$$

   Finally, we have

$$\mathsf{E}[X] \;=\; \int_{-\infty}^{\infty} x f_X(x)\,dx \;=\; \int_0^{8^{1/4}} \frac{\sqrt{8}\,x^2 - x^4}{2}\,dx \;=\; \frac{8^{5/4}}{15} \;=\; 0.897. \quad \diamondsuit$$

4. A tool crib has exponential interarrival and service times, and it serves a very large group of mechanics. The mean time between arrivals is 3 minutes. It takes 2 minutes on average for the single tool-crib attendant to serve a mechanic. What is L, the long-run number of customers (mechanics) in the system?

   Solution: Model the crib as an $M/M/c$ queue with $\lambda = 1/3$, $\mu = 1/2$, and $c = 1$. We have traffic intensity $\rho = \lambda/\mu = 2/3$, and so the $M/M/1$ queueing table gives

$$L \;=\; \frac{\rho}{1-\rho} \;=\; 2. \quad \diamondsuit$$

5. Same as Question #4, except assume that there are now *two* tool-crib attendants.

   Solution: Model the crib as an $M/M/c$ queue with $\lambda = 1/3$, $\mu = 1/2$, and $c = 2$. For this case $c = 2$, we have $\rho = \lambda/(c\mu) = 1/3$, and so the $M/M/2$ queueing table gives

$$P_0 \;=\; \left\{ \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} \right] + \left[\frac{(c\rho)^c}{(c!)(1-\rho)} \right] \right\}^{-1} \;=\; 0.5,$$

   and then

$$L \;=\; c\rho + \frac{(c\rho)^{c+1} P_0}{c(c!)(1-\rho)^2} \;=\; 0.75. \quad \diamondsuit$$

6. Consider Questions #4 and #5. Suppose that an attendant is paid \$8 per hour and mechanics are paid \$10 per hour. Would it be advisable to have one or two tool-crib attendants?

   Solution: The expected cost per hour to run this system is $8c + 10L$.

   For the case $c = 1$, we have

$$8c + 10L \;=\; 8(1) + 10(2) \;=\; 28.$$

   For the case $c = 2$, we have

$$8c + 10L \;=\; 8(2) + 10(0.75) \;=\; 23.5.$$

   So we take $c = 2$. $\diamondsuit$

7. TRUE or FALSE? The runs up-and-down test is used to see if i.i.d. data follow a certain distribution.

   Solution: FALSE. It's an independence test. $\diamondsuit$

8. Suppose that we conduct a Kolmogorov-Smirnov test for a Unif(0,1) distribution and that we observe the three numbers $U_1 = 0.75$, $U_2 = 0.21$, and $U_3 = 0.57$. Find the K-S statistic D. (You do not have to perform the test itself.)

   Solution: Here's our usual table.

$R_{(i)}$	0.210	0.570	0.750
$\frac{i}{n}$	0.333	0.667	1
$\frac{i-1}{n}$	0	0.333	0.667
$\frac{i}{n} - R_{(i)}$	0.123	0.097	0.250
$R_{(i)} - \frac{i-1}{n}$	0.210	0.237	0.083

   Then $D^+ = \max_i\{\frac{i}{n} - R_{(i)}\} = 0.250$, $D^- = \max_i\{R_{(i)} - \frac{i-1}{n}\} = 0.237$, and finally, $D = \max(D^+, D^{-1}) = 0.250$. $\diamondsuit$

9. Find the mean of the random variable generated by the Arena function `DISC(0.3,3,0.8,5,1,10)`.

   Solution: $3(0.3) + 5(0.5) + 10(0.2) = 5.4$. $\diamondsuit$

10. TRUE or FALSE? If you `SEIZE` a member of an Arena resource set, it is possible to `RELEASE` that *specific* member of the set later on.

    Solution: TRUE. $\diamondsuit$

11. The c.d.f. of the Cauchy distribution is $F(x) = \frac{1}{2} + \frac{\tan^{-1}(x)}{\pi}$ for $-\infty < x < \infty$. Write out the appropriate inverse transform equation to generate a Cauchy RV from a Unif(0,1).

Solution: Set $F(X) = \frac{1}{2} + \frac{\tan^{-1}(X)}{\pi} = U$ and solve for X to obtain

$$X \;=\; \tan(\pi(U - 0.5)). \quad \diamondsuit$$

12. Demonstrate your Cauchy generator from Question #11 by using it on $U = 0.95$.

Solution:

$$X \;=\; \tan(\pi(0.95 - 0.5)) \;=\; 6.314. \quad \diamondsuit$$

13. Consider a first-order moving average process, $X_i = \epsilon_i + \theta\epsilon_{i-1}$, $i \geq 1$, where the ϵ_i's are i.i.d. standard normal. Find the lag-1 covariance, $R_1 = \mathsf{Cov}(X_{i+1}, X_i)$.

Solution: From class notes (and it's easy to show anyway), $R_1 = \theta$. $\diamondsuit$

14. Consider the 2×2 covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

Find the Cholesky matrix C such that $\Sigma = CC'$.

Solution: This might be a bit of overkill, but the following algorithm (from class) computes the $k \times k$ matrix C:

For $i = 1, \ldots, k$,

  For $j = 1, \ldots, i-1$,

    $c_{ij} \leftarrow \left(\sigma_{ij} - \sum_{\ell=1}^{j-1} c_{i\ell}c_{j\ell}\right) / c_{jj}$

    $c_{ji} \leftarrow 0$

  $c_{ii} \leftarrow \left(\sigma_{ii} - \sum_{\ell=1}^{i-1} c_{i\ell}^2\right)^{1/2}$

We simply need to substitute $k = 2$, $\sigma_{11} = \sigma_{22} = 1$, and $\sigma_{12} = 0.5$ and then run the algorithm.

For the case $i = 1$, the "For $j = 1, \ldots, i-1$" loop becomes "For $j = 1, 0$", and so does not execute. Thus, for this case, we skip down to the last assignment and set

$$c_{11} \leftarrow \left(\sigma_{11} - \sum_{\ell=1}^{0} c_{1\ell}^2 \right)^{1/2} \;=\; \sigma_{11}^{1/2} \;=\; 1.$$

With this result in hand, we can go to the $i = 2$ case. Now the "For $j = 1, \ldots, i-1$" loop becomes "For $j = 1$", and we can proceed with the following assignments:

$$\begin{aligned} c_{ij} &= c_{21} \leftarrow \left(\sigma_{21} - \textstyle\sum_{\ell=1}^{0} c_{i\ell} c_{1\ell} \right) / c_{11} = \sigma_{21}/c_{11} = 0.5 \\ c_{ji} &= c_{12} \leftarrow 0 \end{aligned}$$

and then the last assignment:

$$c_{ii} \;=\; c_{22} \leftarrow \left(\sigma_{22} - \sum_{\ell=1}^{1} c_{2\ell}^2 \right)^{1/2} \;=\; \left(\sigma_{22} - c_{21}^2 \right)^{1/2} \;=\; (1 - 0.25)^{1/2} = 0.866.$$

So after all of this, here's our final matrix:

$$C \;=\; \begin{pmatrix} 1 & 0 \\ 0.5 & 0.866 \end{pmatrix}.$$

(As a check, you'll see that $CC' = \Sigma$, as desired.) $\diamondsuit$

Alternatively, if you remembered from class notes that C has to be a lower triangular matrix, then you could have solved

$$CC' \;=\; \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ 0 & c_{22} \end{pmatrix} \;=\; \begin{pmatrix} c_{11}^2 & c_{11}c_{21} \\ c_{11}c_{21} & c_{21}^2 c_{22}^2 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \;=\; \Sigma$$

to obtain the same answer as before. $\diamondsuit$

15. What is the name of the method that one usually uses to generate RV's from a nonhomogeneous Poisson process?

    (a) thinning

    (b) independent increment sampling

    (c) antithetic variates

(d) exponential smoothing

Solution: (a) thinning. $\diamondsuit$

16. Suppose that $X(t)$, $t \geq 0$, is a stochastic process with the following properties: (i) $X(0) = 0$, (ii) $X(t) \sim \text{Nor}(0, t)$, and (iii) $X(t)$ has stationary and independent increments. What is the colorful name of the $X(t)$ process?

    Solution: Brownian motion. $\diamondsuit$

17. Find the sample variance of 0, 5, and 10.

    Solution: $S^2 = 25$. $\diamondsuit$

18. If $X_1, \ldots, X_{10}$ are i.i.d. Pois($\lambda = 3$), what is the expected value of the sample variance S^2?

    Solution: $\mathsf{E}[S^2] = \mathsf{Var}(X_i) = 3$. $\diamondsuit$

19. If $X_1, \ldots, X_{10}$ are i.i.d. Pois($\lambda = 3$), what is the expected value of the maximum likelihood estimator for the variance $\mathsf{Var}(X_i)$?

    Solution: By the above problem, $\mathsf{E}[\widehat{\sigma^2}] = \frac{n-1}{n}\mathsf{E}[S^2] = \frac{9}{10}3 = 2.7$. $\diamondsuit$

20. Suppose we observe the Exp(λ) realizations $X_1 = 1.0$, $X_2 = 3.2$, $X_3 = 0.5$, and $X_4 = 1.3$. What is the maximum likelihood estimate of λ?

    Solution: By the class notes, $\hat{\lambda} = 1/\bar{X} = 2/3$. $\diamondsuit$

21. Consider the set-up of Question #20, where $X \sim \text{Exp}(\lambda)$. What is the MLE of $\mathsf{P}(X > 2)$?

Solution: By invariance and the previous problem,

$$\widehat{\mathsf{P}}(X > x) \;=\; e^{-\hat{\lambda}x} \;=\; e^{-(2/3)2} \;=\; e^{-4/3} \;=\; 0.264. \quad \diamondsuit$$

22. TRUE or FALSE? The MLE's for the gamma distribution require an evaluation of the *digamma* function.

    Solution: TRUE. $\diamondsuit$

23. Find the MLE for θ if X has p.d.f.

$$f(x) \;=\; (\theta+1)x^{\theta}, \quad 0 \le x \le 1, \quad \text{where } \theta > 0.$$

Solution: The likelihood function is

$$L(\theta) \;=\; \prod_{i=1}^{n} f(x_i) \;=\; (\theta+1)^n \prod_{i=1}^{n} x_i^{\theta}.$$

Thus,

$$\ell\text{n}(L(\theta)) \;=\; n\,\ell\text{n}(\theta+1) + \theta\,\ell\text{n}\left(\prod_{i=1}^{n} x_i\right).$$

This implies that

$$\frac{d}{d\theta}\,\ell\text{n}(L(\theta)) \;=\; \frac{n}{\theta+1} + \ell\text{n}\left(\prod_{i=1}^{n} x_i\right).$$

Setting the derivative to 0 and solving yields

$$\hat{\theta} \;=\; \frac{-n}{\ell\text{n}(\prod_{i=1}^{n} x_i)} - 1. \quad \diamondsuit$$

24. Suppose we obtain sample data $X_1 = 0.30$, $X_2 = 0.78$, and $X_3 = 0.11$ from the distribution described in Question #23. What is the value of the MLE?

Solution:

$$\hat{\theta} \;=\; \frac{-n}{\ell\text{n}(\prod_{i=1}^{n} x_i)} - 1 \;=\; -0.180. \quad \diamondsuit$$

25. Suppose we're conducting a χ^2 goodness-of-fit test to determine whether or not 50 i.i.d. observations are from the distribution described in Question #23. If we divide the observations into 5 equal-probability intervals, how many degrees of freedom will our test have?

    Solution: We have $k = 5$ intervals and $s = 1$ unknown parameter. Thus, $\nu = k - 1 - s = 3$. $\diamondsuit$

26. Generally speaking, does output analysis for terminating simulation most-often use the method of independent replications or the method of batch means?

    Solution: IR. $\diamondsuit$

27. Consider the following 9 observations arising from a simulation:

    183 154 180 175 162 200 173 191 183

    Use the method of batch means to calculate a two-sided 95% confidence interval for the mean μ. In particular, use three batches of size three.

    Solution: The batch size is $m = 3$, the number of batches is $b = 3$, and the total number of observations is $n = 9$. The grand sample mean is $\bar{Y}_9 = 177.89$. The batch means are

    $$\bar{Y}_{1,3} = 172.33, \quad \bar{Y}_{2,3} = 179.00, \quad \bar{Y}_{3,3} = 182.33.$$

    The batch means variance estimator is

    $$\widehat{V}_B \;=\; \frac{m}{b-1}\sum_{i=1}^{b}(\bar{Y}_{i,3} - \bar{Y}_n)^2 \;=\; 77.79.$$

    The batch means confidence interval is

    $$\begin{aligned} \mu &\in \bar{Y}_n \pm t_{\alpha/2,b-1}\sqrt{\widehat{V}_B/n} \\ &= 177.89 \pm t_{0.025,2}\sqrt{77.79/9} \\ &= 177.89 \pm 4.30(2.94) \\ &= 177.89 \pm 12.64 \;=\; [165.25,\ 190.53]. \quad \diamondsuit \end{aligned}$$

28. Suppose $[-1, 1]$ is a 95% nonoverlapping batch means confidence interval for the mean μ based on 5 batches of size 500. Now the boss has decided that she wants a 90% CI based on those same 5 batches of size 500. What is it?

Solution: The confidence interval is of the form

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

This implies that $\bar{Y}_n = 0$ and the half-length is $t_{0.025,4}\sqrt{\widehat{V}_B/n} = 1$. Thus, the new 90% confidence interval is

$$\begin{aligned} \text{new CI} &= \bar{Y}_n \pm t_{0.05,4}\sqrt{\widehat{V}_B/n} \\ &= \pm\frac{t_{0.05,4}}{t_{0.025,4}} t_{0.025,4}\sqrt{\widehat{V}_B/n} \\ &= \pm\frac{2.13}{2.78} \times 1 \\ &= \pm 0.766. \quad \diamondsuit \end{aligned}$$

29. Consider the output analysis method of nonoverlapping batch means. Assuming that you have a sufficiently large batch size, it can be shown that when the number of batches b is *even*, the *expected width* of the 95% two-sided confidence interval for μ is proportional to

$$\frac{t_{0.025,b-1}}{\sqrt{b-1}} \frac{\left(\frac{b-1}{2}\right)\left(\frac{b-3}{2}\right)\cdots\frac{1}{2}}{\left(\frac{b-2}{2}\right)!}.$$

Using the above equation, determine which of $b = 4$ or $b = 6$ or $b = 8$ gives the smallest expected width.

Solution: Let $h(b)$ denote the value of the above expression as a function of b. Then easy calculations reveal that $h(4) = 1.377$, $h(6) = 1.078$, and $h(8) = 0.976$. So the answer is $b = 8$. $\diamondsuit$

30. Consider the 9 observations from Question #27. If we choose a batch size of 7, calculate all of the *overlapping* batch means for me.

Solution:

$\bar{Y}^{\text{o}}_{1,7} = \frac{1}{7}\sum_{i=1}^{7} Y_i = 175.29$,

$\bar{Y}^{\text{o}}_{2,7} = \frac{1}{7}\sum_{i=2}^{8} Y_i = 176.43$, and
$\bar{Y}^{\text{o}}_{3,7} = \frac{1}{7}\sum_{i=3}^{9} Y_i = 180.57$. $\diamondsuit$

31. Consider a particular data set of 30000 stationary waiting times obtained from a large queueing system. Suppose we choose a batch size of $m = 1000$. Approximately how many degrees of freedom would the corresponding *overlapping* batch means confidence interval have?

Solution: Define the quantity $b = n/m = 30$. From class notes, the approximate d.f. is about $\frac{3b}{2} = 45$. $\diamondsuit$

32. Suppose that $X_1, X_2, \ldots$ is a stationary stochastic process with covariance function $R_0 = 3$, $R_{\pm 1} = 2$, $R_{\pm 2} = 1$, and $R_k = 0$ for all other values of k. If $\bar{X}_n = \frac{1}{10}\sum_{i=1}^{10} X_i$ is the sample mean based on the first 10 observations, find $\mathsf{Var}(\bar{X}_n)$.

Solution: From class, we know that

$$\begin{aligned}
\mathsf{Var}(\bar{X}_n) &= \frac{1}{n}\left[R_0 + 2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)R_k\right] \\
&= \frac{1}{10}\left[R_0 + 2\sum_{k=1}^{9}\left(1-\frac{k}{10}\right)R_k\right] \\
&= \frac{1}{10}\left[3 + 2\left(1-\frac{1}{10}\right)(2) + 2\left(1-\frac{2}{10}\right)(1)\right] \\
&= 0.82 \quad \diamondsuit
\end{aligned}$$

33. Suppose that I'm using simulation to compare the waiting-time results from two alternative configurations of a certain service system. For Configuration A, I simulate 4 independent replications, and I'm willing to assume that the following replicate means are approximately i.i.d. normal:

22.3 6.4 11.4 31.6

These answers are bouncing all over the place, eh? I also run 4 independent replications of Configuration B, though I try to make an apples-to-apples comparison with the runs from Configuration A by keeping the same arrivals and service times. The corresponding results for Configuration B are:

32.3 15.9 21.9 41.2

These runs also bounce around a lot. In any case, I'd like you to give me a two-sided 95% paired-t confidence interval for the difference in the means of Configurations A and B.

Solution: The CI is based on the differences between corresponding observations, $D_1 = -10.0$, $D_2 = -9.5$, $D_3 = -10.5$, and $D_4 = -9.6$.

$$\begin{aligned}
\mu_A - \mu_B &\in \bar{D} \pm t_{\alpha/2,n-1}\sqrt{S_D^2/n} \\
&= -9.9 \pm t_{0.025,3}\sqrt{0.207/4} \\
&= -9.9 \pm 3.18(0.227) \\
&= -9.9 \pm 0.72 \;=\; [-10.62,\, -9.18]. \quad \diamondsuit
\end{aligned}$$

34. Let's estimate the mean of an exponential distribution via MC simulation. As we know, you can write an Exp(1) RV as $X = -\ell\mathrm{n}(1-U)$ using inversion. Similarly, we know that the antithetic version, $Y = -\ell\mathrm{n}(U)$, is also Exp(1). As we explained in class, X and Y are negatively correlated; and this fact can be put to good use in estimating the mean. In order to do this, consider the following 4 Unif(0,1) PRN's:

$$0.18 \quad 0.77 \quad 0.83 \quad 0.55$$

I want you to calculate and report 3 estimators for the mean: $\bar{X}$, $\bar{Y}$, and $\bar{Z} = (\bar{X}+\bar{Y})/2$. Also, which is the best? (Note that plenty of room has been provided on the answer sheet for you to report these 4 items.)

Solution: After some easy algebra, we find that

$$\bar{X} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i \;=\; \frac{-1}{n}\sum_{i=1}^{n} \ell\mathrm{n}(1-U_i) \;=\; \frac{-1}{4}\,\ell\mathrm{n}\left[\prod_{i=1}^{4}(1-U_i)\right] \;=\; 1.0596. \quad \diamondsuit$$

Similarly,

$$\bar{Y} \;=\; \frac{1}{n}\sum_{i=1}^{n} Y_i \;=\; \frac{-1}{4}\,\ell\mathrm{n}\left[\prod_{i=1}^{4} U_i\right] \;=\; 0.6901. \quad \diamondsuit$$

And then we have

$$\bar{Z} \;=\; \frac{\bar{X}+\bar{Y}}{2} \;=\; 0.8749. \quad \diamondsuit$$

Of course, the actual mean of the Exp(1) is 1. All of the estimators are unbiased, but estimator $\bar{X}$ comes closest to the right answer in this instance. This is just bad luck, since estimator $\bar{Z}$ has the smallest variance — it will usually win, but not this time. $\diamondsuit$

Table 1: Standard normal values

z	1	1.28	1.5	1.645	1.96	2
$\mathsf{P}(Z \leq z)$	0.8413	0.9000	0.9332	0.9500	0.9750	0.9773

Table 2: $\chi^2_{\alpha,\nu}$ values

$\nu \setminus \alpha$	0.10	0.05	0.025
1	2.71	3.84	5.02
2	4.61	5.99	7.38
3	6.25	7.81	9.35
4	7.78	9.49	11.14
5	9.24	11.07	12.83
6	10.65	12.59	14.45

Table 3: $t_{\alpha,\nu}$ values

$\nu \setminus \alpha$	0.10	0.05	0.025
1	3.08	6.31	12.71
2	1.89	2.92	4.30
3	1.64	2.35	3.18
4	1.53	2.13	2.78
5	1.48	2.02	2.57
6	1.44	1.94	2.45
7	1.42	1.90	2.36
8	1.40	1.86	2.31
9	1.38	1.83	2.26
10	1.37	1.81	2.23

Table 4: $F_{0.025,m,n}$ values

$n \setminus m$	3	4	5
3	15.44	15.10	14.88
4	9.98	9.60	9.36
5	7.76	7.39	7.15