**NAME** $\rightarrow$

# ISyE 3044 — Fall 2013 — Test #3 Solutions

(revised 4/22/14)

This test 2 hours. You are allowed three cheat sheets. Every question is 3 points. *Only turn in succinct answers.* Good luck!

1. If $X$ and $Y$ have joint p.d.f. $f(x,y) = c(1-x)y$, $0 \le y \le x \le 1$, for some appropriate constant $c$, find $\mathsf{E}[X]$.

   **Solution:** First of all,

$$1 \;=\; \int_0^1 \int_0^x c(1-x)y\,dy\,dx \;=\; \frac{c}{2}\int_0^1 (x^2-x^3)\,dx \;=\; \frac{c}{24},$$

   so that $c = 24$. Then

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy \\ &= \int_0^x 24(1-x)y\,dy \\ &= 12(x^2-x^3), \quad 0<x<1. \end{aligned}$$

   Thus,

$$\mathsf{E}[X] \;=\; \int_0^1 12(x^3-x^4)\,dx \;=\; 0.6. \quad \Box$$

2. If $X$ and $Y$ are i.i.d. Pois(3) RV's, find $\mathsf{Var}(XY)$.

   **Solution:**

$$\begin{aligned} \mathsf{Var}(XY) &= \mathsf{E}[(XY)^2] - (\mathsf{E}[XY])^2 \\ &= \mathsf{E}[X^2]\mathsf{E}[Y^2] - (\mathsf{E}[X]\mathsf{E}[Y])^2 \quad X,Y \text{ indep} \\ &= (\mathsf{E}[X^2])^2 - (\mathsf{E}[X])^4 \quad X,Y \text{ i.i.d.} \\ &= (\mathsf{Var}(X) + (\mathsf{E}[X])^2)^2 - (\mathsf{E}[X])^4 \quad X,Y \text{ i.i.d.} \\ &= (3+9)^2 - 3^4 \;=\; 63. \quad \Box \end{aligned}$$

3. Suppose that customers to a single-server queue according to a Poisson process with rate 5/hour. Service times are i.i.d. exponential with a mean of 15 minutes. It turns out that the system has a capacity of 4 (3 in line and one being served). What is the server utilization (percent of time that the server is being used)?

   **Solution:** This is an M/M/1/4 queuing system. From the tables with $a = \lambda/\mu = 5/4$, we have

   $$\rho = \frac{\lambda_e}{\mu} = 1 - P_0 = 1 - \frac{(1-a)a^0}{1-a^{N+1}} = 0.8782. \quad \Box$$

4. Draw an Arena `DECIDE` block.

   **Solution:** It's a sideways diamond. $\Box$

5. In what Arena template would you find the `Sequence` spreadsheet?

   **Solution:** `Advanced Transfer`. $\Box$

6. In Arena, you can `LEAVE` a station. What is the analogous block to use to get into a station?

   **Solution:** `ENTER`. $\Box$

7. Which method is *not* a viable way for an entity to LEAVE a resource? (i) Route; (ii) Connect; (iii) Move; (iv) Transport; (v) Convey. (There may be more than one correct answer.)

   **Solution:** (iii) Move. $\Box$

8. TRUE or FALSE? Suppose that $U_1, U_2, \ldots$ are truly i.i.d. U(0,1) random variables. Then a 95% chi-square goodness-of-fit test for uniformity will *incorrectly* reject uniformity of the observations about 5% of the time.

   **Solution:** TRUE. (That's what Type I error is.) $\Box$

9. TRUE or FALSE? Suppose $Z_1, \ldots, Z_6$ are i.i.d. standard normal random variables obtained by the Box–Muller method. Then $\sum_{i=1}^{6} Z_i^2 \sim \text{Erlang}_3(1/2)$.

   **Solution:** TRUE. The sum of 6 i.i.d. $Z_i^2$ random variables is a $\chi^2(6)$, which is itself an $\text{Erlang}_3(1/2)$. $\square$

10. Suppose $X_1, X_2, \ldots, X_{100}$ are i.i.d. with $\Pr(X = 0) = \Pr(X = 2) = 0.5$. Define the sample mean $\bar{X} \equiv \sum_{i=1}^{100} X_i/100$. Use the Central Limit Theorem to find an approximate expression for $\Pr(\bar{X} < 1.1)$.

    **Solution:** Note that $\mathsf{E}[X] = \sum_x x \Pr(X = x) = 1$ and $\mathsf{E}[X^2] = \sum_x x^2 \Pr(X = x) = 2$, so that $\mathsf{Var}(X) = \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = 1$. These results makes sense since $X \sim 2\text{Bern}(0.5)$.

    Anyway, the CLT now implies that $\bar{X} \approx \mathsf{Nor}(\mu, \sigma^2/n) = \mathsf{Nor}(1, 0.01)$, and so

$$\Pr(\bar{X} < 1.1) \;=\; \Pr\left(\frac{\bar{X} - 1}{\sqrt{0.01}} < \frac{1.1 - 1}{\sqrt{0.01}}\right) \;\approx\; \Pr(\mathsf{Nor}(0,1) < 1) \;=\; 0.8413. \quad \square$$

11. If $U, V$ are i.i.d. U(0,1), what's the distribution of $-\ell\text{n}(\sqrt{U}) - \ell\text{n}(\sqrt{V})$?

    **Solution:**

$$-\ell\text{n}(\sqrt{U}) - \ell\text{n}(\sqrt{V}) \;=\; -\frac{1}{2}\ell\text{n}(U) - \frac{1}{2}\ell\text{n}(V) \;\sim\; \text{Erlang}_2(2). \quad \square$$

12. Consider the stationary first-order exponential autoregressive process (EAR(1)),

$$X_i \;=\; \begin{cases} \alpha X_{i-1}, & \text{w.p. } \alpha \\ \alpha X_{i-1} + \epsilon_i, & \text{w.p. } 1 - \alpha, \end{cases}$$

    where $X_0$ and the $\epsilon_i$'s are i.i.d. $\text{Exp}(\lambda)$, and $0 < \alpha < 1$. Find $\mathsf{Cov}(X_0, X_1)$.

    **Solution:**

$$\begin{aligned} \mathsf{Cov}(X_0, X_1) &= \mathsf{Cov}\Big(X_0,\, \alpha X_0\Big)\alpha + \mathsf{Cov}\Big(X_0,\, \alpha X_0 + \epsilon_1\Big)(1 - \alpha) \\ &= \alpha^2 \mathsf{Var}(X_0) + \alpha(1 - \alpha)\mathsf{Var}(X_0) + 0 \\ &= \alpha/\lambda^2. \quad \square \end{aligned}$$

13. If $X_1, \ldots, X_n$ are i.i.d. Exp(1/9), what is the expected value of the sample variance $S^2$?

    **Solution:** $\mathsf{E}[S^2] = \sigma^2 = 81.$ $\square$

14. If $X_1, X_2, X_3$ are i.i.d. normal, with $X_1 = 3$, $X_2 = 2$, and $X_3 = 7$, what is the MLE for $\mathsf{E}[X_i^2]$?

    **Solution:** By invariance, the MLE is

$$\widehat{\mathsf{E}}[X_i^2] \;=\; \hat{\mu}^2 + \hat{\sigma}^2 \;=\; \bar{X}^2 + \frac{n-1}{n} S^2 \;=\; 4^2 + \frac{14}{3} \;=\; 20.67. \quad \square$$

15. Find $x$ such that $e^{2x} = 1/x$. (Get within two decimals.)

    **Solution:** Bisection search quickly reveals that $x$ is about 0.4265. $\square$

16. TRUE or FALSE? The square root of the sample variance is unbiased for the standard deviation.

    **Solution:** FALSE. ($\mathsf{E}[S^2] = \sigma^2 \not\Rightarrow \mathsf{E}[S] = \sigma$.) $\square$

17. Suppose we're conducting a $\chi^2$ goodness-of-fit test to determine whether or not 200 i.i.d. observations are from a Johnson distribution, which has 4 parameters that must be estimated. If we divide the observations into 10 equal-probability intervals, how many degrees of freedom will our test have?

    **Solution:** $10 - 4 - 1 = 5.$ $\square$

18. We are interested in seeing if the number of emergency department visits occurring each day at the Georgia Tech clinic is Binomial(5,0.5). Below are the results for a 200-day period. We'll assume that the numbers from day to day are i.i.d.

| # of visits | # of days |
|---|---|
| 0 | 6 |
| 1 | 30 |
| 2 | 61 |
| 3 | 62 |
| 4 | 34 |
| 5 | 7 |

Thus, for example, there were 30 days during when the ED had exactly 1 visit.

We'll perform a 95% $\chi^2$ goodness-of-fit test to see if the number of accidents each day is Binomial(5,0.5).

(a) How many intervals will you use for your test?

**Solution:** Let $X$ denote the number of visits on a particular day. Under the null hypothesis, the expected number of occurrences of $i$ visits is $E_i = n\Pr(X = i) = 200\binom{5}{i}(0.5)^5$. So we have the following table.

| $i$ | $O_i$ | $E_i$ |
|---|---|---|
| 0 | 6 | 6.25 |
| 1 | 30 | 31.25 |
| 2 | 61 | 62.5 |
| 3 | 62 | 62.5 |
| 4 | 34 | 31.25 |
| 5 | 7 | 6.25 |

Since all of the $E_i$'s are $\geq 5$, we can use all of the cells. Thus, the answer is 6 intervals. $\Box$

(b) What is your statistic value?

**Solution:** $\chi_0^2 = \sum_{i=0}^{5}(O_i - E_i)^2/E_i = 0.432.$ $\Box$

(c) What is your conclusion? Binomial(5,0.5) or not?

**Solution:** Since the previous answer is so small, we don't really need to look up the quantile, but I'll do it anyway. The critical value is $\chi^2_{\alpha,k-1} = \chi^2_{0.05,5} = 11.07$, indicating a fail to reject. (So, yes, we'll assume it's Binomial.) $\Box$

19. Consider the following PRN's: 0.18, 0.92, 0.61, 0.33. If we use the Kolmogorov-Smirnov goodness-of-fit test to see if these numbers are U(0,1), what is the value of the test statistic?

**Solution:** Let's make the usual table with the ordered PRN's.

| $i$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $R_{(i)}$ | 0.18 | 0.33 | 0.61 | 0.92 |
| $\frac{i}{n} - R_{(i)}$ | 0.07 | 0.17 | 0.14 | 0.07 |
| $R_{(i)} - \frac{i-1}{n}$ | 0.18 | 0.08 | 0.11 | 0.18 |

This indicates that $D_n^+ = \max_i[\frac{i}{n} - R_{(i)}] = 0.17$ and $D_n^- = \max_i[R_{(i)} - \frac{i-1}{n}] = 0.18$, so that $D_n = \max(D^+, D^-) = 0.18$. $\Box$

20. Do the PRN's in Question 19 pass the Kolmogorov–Smirnov goodness-of-fit test for uniformity at level $\alpha = 0.05$?

**Solution:** From the one-sided table, we have $D_{\alpha,n} = D_{0.05,4} = 0.565$. Since $D_n < D_{\alpha,n}$, we fail to reject uniformity. $\Box$

21. Consider a stationary stochastic process $X_1, X_2, \ldots,$ with covariance function $R_k = \mathsf{Cov}(X_1, X_{1+k}) = 3 - k$ for $k = 0, 1, 2, 3$, and $R_k = 0$ for $k \geq 4$. Find $\mathsf{Var}(\bar{X}_4)$.

**Solution:** From class notes, we have

$$
\begin{aligned}
\mathsf{Var}(\bar{X}_n) &= \frac{1}{n}\left[R_0 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)R_k\right] \\
&= \frac{1}{4}\left[R_0 + 2(\frac{3}{4})R_1 + 2(\frac{2}{4})R_2 + +2(\frac{1}{4})R_3\right] \\
&= \frac{1}{4}\left[3 + 2(\frac{3}{4})2 + 2(\frac{2}{4})1 + +2(\frac{1}{4})0\right] \\
&= 1.75. \quad \Box
\end{aligned}
$$

22. Consider the following (approximately normal) average waiting times from 4 independent replications of a complicated queueing network. Suppose that each output is based on the average of 500 waiting times:

30 40 10 50

Use the method of independent replications to calculate a two-sided 90% confidence interval for the mean $\mu$.

**Solution:** We use $b = 4$ reps here. $\bar{Z}_b = \sum_{i=1}^{b} Z_i/b = 32.5$ and $S_Z^2 = \frac{1}{b-1}\sum_{i=1}^{b}(Z_i - \bar{Z}_b)^2 = 291.67$. Then the desired CI is

$$\begin{aligned} \mu &\in \bar{Z}_b \pm t_{\alpha/2,b-1}\sqrt{S_Z^2/b} \\ &= 32.5 \pm t_{0.05,3}\sqrt{291.67/4} \\ &= 32.5 \pm 2.353(8.539) \\ &= 32.5 \pm 20.09 \;=\; [12.4,\, 52.6]. \quad \Box \end{aligned}$$

23. Which is the method of batch means more appropriate for: terminating or steady-state simulations?

    **Solution:** Steady-State. $\Box$

24. Which is usually a better way to deal with initialization bias in steady-state simulation analysis: (i) make an extremely long run to overwhelm the bias, or (ii) perform truncation?

    **Solution:** Truncate. $\Box$

25. Consider the following 10 snowfall totals in Buffoonalo, NY over consecutive years:

    130 94 125 112 150 123 141 133 128 152

    Use the method of batch means to calculate a two-sided 90% confidence interval for the mean $\mu$. In particular, use two batches of size 5.

    **Solution:** We use $b = 2$ batches here here. $\bar{X}_n = 128.8$, the batch means are $\bar{X}_{1,5} = 122.2$ and $\bar{X}_{2,5} = 135.4$, and the batch means estimator for the variance parameter is

$$\widehat{V}_B \;=\; \frac{m}{b-1}\sum_{i=1}^{b}(\bar{X}_{i,m} - \bar{X}_n)^2 \;=\; 5\sum_{i=1}^{2}(\bar{X}_{i,5} - \bar{X}_{10})^2 \;=\; 435.6.$$

Then the desired CI is

$$\begin{aligned}\mu &\in \bar{X}_n \pm t_{\alpha/2,b-1}\sqrt{\widehat{V}_B/n} \\ &= 128.8 \pm t_{0.05,1}\sqrt{435.6/10} \\ &= 128.8 \pm 6.314(6.6) \\ &= 128.8 \pm 41.67 \;=\; [87.1,\, 170.5]. \quad \Box\end{aligned}$$

26. Suppose $[0,1]$ is a 90% confidence interval for the mean $\mu$ based on 10 independent replications of size 1000. Now the boss has decided that she wants a 99% CI for $\mathbf{2}\mu$ based on those same 10 replications of size 1000. What is it?

**Solution:** Let's get a 99% CI for $\mu$ first. To begin with, the original 90% CI for $\mu$ is

$$\mu \in \bar{Z}_b \pm t_{0.05,\,b-1}\sqrt{\frac{S_Z^2}{b}} = 0.5 \pm 0.5.$$

This implies that the 99% CI for $\mu$ will have half-length

$$t_{0.005,\,b-1}\sqrt{\frac{S_Z^2}{b}} \;=\; \frac{t_{0.005,\,b-1}}{t_{0.05,\,b-1}} t_{0.05,\,b-1}\sqrt{\frac{S_Z^2}{b}} \;=\; \frac{t_{0.005,\,9}}{t_{0.05,\,9}}(0.5) \;=\; \frac{3.250}{1.833}(0.5) \;=\; 0.887.$$

Now we can get the 99% CI for $\mu$:

$$\mu \in \bar{Z}_b \pm t_{0.005,\,b-1}\sqrt{\frac{S_Z^2}{b}} \;=\; 0.5 \pm 0.887.$$

This immediately implies that the 99% CI for $2\mu$ is

$$2\mu \in 1 \pm 1.77 \;=\; [-0.77,\, 2.77]. \quad \Box$$

27. Suppose I use the method of overlapping batch means with sample size $n = 10000$ and batch size $m = 500$. Approximately how many degrees of freedom will the resulting variance estimator have?

**Solution:** Denote $b = n/m = 20$. You get approximately $\frac{3}{2}(b-1) = 28.5$ d.f. (Will also accept $3b/2 = 30$ or anything reasonably close.) $\Box$

28. If $\mathcal{W}(t)$ is a standard Brownian motion process. What is the distribution of $\mathcal{W}(1)+\mathcal{W}(2)$?

**Solution:** Note that

$$\mathsf{Var}\Big(\mathcal{W}(1)+\mathcal{W}(2)\Big) \;=\; \mathsf{Var}(\mathcal{W}(1))+\mathsf{Var}(\mathcal{W}(2))+2\mathsf{Cov}\Big(\mathcal{W}(1),\,\mathcal{W}(2)\Big) \;=\; 1+2+2 \;=\; 5.$$

Thus, $\mathcal{W}(1)+\mathcal{W}(2) \sim \mathsf{Nor}(0,5)$. □

29. Let's use the basic Monte Carlo technique from class to integrate $I = \int_0^1 e^x\,dx$.

(a) First of all, what is the exact value of $I$?

**Solution:** $I = e - 1 = 1.718$. □

(b) Use the PRN's 0.95, 0.63, 0.15, and 0.42 to estimate $I$.

**Solution:** $\hat{I}_n = \frac{1}{n}\sum_{i=1}^n e^{U_i} = 1.787$. □

(c) Now use antithetics to estimate $I$.

**Solution:** $\tilde{I}_n = \frac{1}{n}\sum_{i=1}^n e^{1-U_i} = 1.656$. □

(d) Combine your last two answers.

**Solution:** $\bar{I}_n = \frac{1}{2}(\hat{I}_n + \tilde{I}_n) = 1.721$, which is a great answer. □