Name: ----------

Directions: **Work only on this sheet** (on both sides, if needed); do not turn in any supplementary sheets of paper. There is actually plenty of room for your answers, as long as you organize yourself BEFORE starting writing. In order to get full credit, **SHOW YOUR WORK**.

IMPORTANT NOTE: Leave definite integrals as is. For example, write $\int_0^1 x^2\ dx$ instead of 1/3.

1. Consider the bus example in Section 5 of the PLN which reviews continuous random variables.

(a) () Find $P(Y > 10)$.

(b) () Write $Y = Y_1 + Y_2$, with Y_1 being the time since the last bus when I arrive, and with Y_2 being the time that I wait for the next bus. The fact that Y has an Erlang distribution tells us what about Y_1 and Y_2?

2. () In our ordinary coins which we use every day, each one has a slightly different probability of heads, R. Say R has the distribution $N(0.5, 0.03^2)$. Choose a coin from a batch at random, then toss it 10 times. Let N be the number of heads we get. Find Var(N).

3. Say we roll a die n times. Let $C_n = (N_{1n}, ..., N_{5n})$, where N_{in} is the number of rolls for which the side facing upward shows i dots.

(a) Why is C_n approximately normal?

(b) What is its covariance matrix?

Solutions:

1.a $\int_{10}^{\infty} 0.01te^{-0.1t}\ dt$

1.b As noted in the PLN, Y has a two-component Erlang distribution, so it is a sum of two independent exponentially distributed random variables with the same parameter. By the Markov property, Y_2 is exponentially distributed, so Y_1 must be the "other" exponential component. So, we can conclude that (i) Y_1 and Y_2 are exponentially distributed with mean 10, (ii) they are independent.

2. From the PLN, we know that

$$Var(N) = E[Var(N|R)] + Var[E(N|R)] \tag{1}$$

Note carefully that even though there are integrals involved here, there will be **NO EXPLICIT INTEGRALS IN THIS SOLUTION.** This illustrates the real usefulness of the Theorem of Total Expectation.

Let's evaluate E[Var(N|R)] first. The conditional distribution of N, given R, is binomial with n = 10 and p = R. Therefore

$$Var(N|R) = 10R(1-R) \tag{2}$$

and thus

$$E[Var(N|R)] = 10\left[E(R) - E(R^2)\right] \tag{3}$$

We can get $E(R^2)$ from the relation $Var(R) = E(R^2) - (ER)^2$. Substituting this in (3), we have

$$E[Var(N|R)] = 10\left[0.5 - (0.03^2 + 0.5^2)\right] \approx 2.5 \tag{4}$$

As to the second term in (1), we have that

$$Var[E(N|R)] = Var(10R) = 100 \cdot 0.03^2 \approx 0.1 \tag{5}$$

Thus $Var(N) \approx 2.6$.

3.a This is due to the *multivariate* Central Limit Theorem, applied to the vectors $U_i = (V_{1i}, ..., V_{5i})$, where V_{ji} is 1 if the i^{th} roll of the die produces j, 0 otherwise, j = 1,...,5.

(Note carefully that this is a generalization of the analysis we used in our PLN unit on continuous random variables to show that the binomial is approximately normal.)

Since

$$C_n = \sum_{i=1}^{n} U_i \tag{6}$$

the multivariate CLT says that C_n is approximately normal.

3.b

$$\begin{aligned} Cov(N_{rn}, N_{sn}) &= Cov\left[\sum_{i=1}^{n} V_{ri}, \sum_{k=1}^{n} V_{sk}\right] && (7)\\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} Cov(V_{ri}, V_{sk}) && (8)\\ &= \sum_{i=1}^{n} Cov(V_{ri}, V_{si}) && (9) \end{aligned}$$

where the second equality comes from the bilinearity of $Cov(\cdot,\cdot)$ and the third comes from the fact that the independence of the rolls of the die implies that $Cov(V_{ri}, V_{sk}) = 0$ when $i \neq k$.

Now for $r \neq s$,

$$\begin{aligned} Cov(V_{ri}, V_{si}) &= E(V_{ri}V_{si}) - E(V_{ri})E(V_{si}) && (10)\\ &= 0 - \frac{1}{6}^2 && (11)\\ &= -\frac{1}{36} && (12) \end{aligned}$$

So,

$$Cov(N_{rn}, N_{sn}) = \begin{cases} n \cdot \frac{5}{36}, & \text{if r} = \text{s} \\ -n \cdot \frac{1}{36}, & \text{if r} \neq \text{s} \end{cases} \tag{13}$$

Note by the way that the reason I didn't include a N_{6n} is that it is a linear combination of the other variables.