Math 461 Fall 2006 — Test 2 Solutions

Total points: **100**. Do all questions. **Explain** all answers. No notes, books, or electronic devices.

1. [10=5+5 points] Assume $X \sim$ Exponential(λ). Justify the following two formulas, by directly using the exponential density function.

(a) $P(0 < X < b) = 1 - e^{-\lambda b}$

Solution. The exponential density is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$, and $f(x) = 0$ for $x < 0$. We integrate the density to evaluate the probability:

$$\begin{aligned} P(0 < X < b) &= \int_0^b f(x)\,dx \\ &= \int_0^b \lambda e^{-\lambda x}\,dx \\ &= -e^{-\lambda x}\Big|_0^b \\ &= 1 - e^{-\lambda b}. \end{aligned}$$

(b) $E[X] = 1/\lambda$

Solution.

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x)\,dx \\ &= \int_0^{\infty} x\lambda e^{-\lambda x}\,dx \\ &= \frac{1}{\lambda}\int_0^{\infty} y e^{-y}\,dy && \text{where } y = \lambda x \text{ and } dy = \lambda dx, \\ &= \frac{1}{\lambda}\left(y(-e^{-y})\Big|_0^{\infty} - \int_0^{\infty} (-e^{-y})\,dy \right) && \text{by integration by parts} \\ &= \frac{1}{\lambda}\left(0 + \int_0^{\infty} e^{-y}\,dy \right) \\ &= \frac{1}{\lambda}. \end{aligned}$$

2. [15 points] An item is manufactured so that its width is normally distributed with mean $\mu = 900$ units and standard deviation σ units.

What is the largest allowable value of σ such that at least 99% of the items will have widths in the range from 895 to 905 units?

Solution. Write $X \sim \text{Normal}(900, \sigma^2)$ for the width. We want

$$\begin{aligned}
.99 &\leq P(895 < X < 905) \\
&= P\left(\frac{895-900}{\sigma} < \frac{X-900}{\sigma} < \frac{905-900}{\sigma}\right) \\
&= P\left(-\frac{5}{\sigma} < Z < \frac{5}{\sigma}\right) \qquad \text{where } Z \text{ is standard normal} \\
&= \Phi(5/\sigma) - \Phi(-5/\sigma) \\
&= 2\Phi(5/\sigma) - 1 \qquad \text{(why?)}.
\end{aligned}$$

Rearranging the inequality, we want

$$.995 \leq \Phi(5/\sigma),$$

and after consulting our Table of the standard normal distribution, we see the inequality holds provided $2.58 \leq 5/\sigma$, or $\sigma \leq 5/2.58 \approx 1.94$.

Hence the largest allowable value of σ is 1.94 units.

Remarks.

1. There is no "continuity correction" needed in this problem, because we are not approximating with a Normal random variable — the random variable is already Normal!

2. This was a homework problem.

3. [15 points] Let $X \sim \text{Uniform}(0,1)$. Find the density function of $Y = e^X$.

Solution. First we determine which values Y can take on: X takes values between 0 and 1, and so $Y = e^X$ takes values between $e^0 = 1$ and $e^1 = e$.

For any a and b in that range, $1 < a \le b < e$, we have

$$\begin{aligned}
P(a < Y < b) &= P(a < e^X < b) \\
&= P(\log a < X < \log b) \\
&= \log b - \log a && \text{since } X \text{ is uniform on } (0,1) \\
&= F(b) - F(a) && \text{where } F(y) = \log y \\
&= \int_a^b F'(y)\,dy && \text{by the Fundamental Theorem} \\
&= \int_a^b \frac{1}{y}\,dy && \text{since } F'(y) = 1/y.
\end{aligned}$$

Thus the density of Y is $g(y) = 1/y$ for $1 < y < e$, and $g(y) = 0$ otherwise.

Remarks.

1. This problem was recommended as preparation on the Test 2 handout.
2. The general formula for transforming a density from X to Y is $g(y) = f(x)\left|\frac{dx}{dy}\right|$. In this problem $y = e^x$, and so $x = \log y$ and $\frac{dx}{dy} = \frac{1}{y}$; also we know $f(x) = 1$ if $0 < x = e^y < 1$, and so we arrive at the same answer as above.

But this general formula can be tricky to apply in practice, especially when more than one x value gives the same y value (*e.g.* $y = x^2$). So it is better to argue directly, like we did in the Solution above.

4. [15 points] Suppose $X \sim$ Geometric(p), and $P(X > k) \leq 1/2$. Show $k \geq \log 2/\log(1/(1-p))$.

Solution. Recall X represents the number of trials for the first success, in an infinite sequence of Bernoulli trials. So $X > k$ means precisely that the first k trials are failures. This occurs with probability $(1-p)^k$, and hence

$$\begin{aligned}\frac{1}{2} &\geq P(X > k) \\ &= (1-p)^k.\end{aligned}$$

Rearranging the inequality gives

$$\left(\frac{1}{1-p}\right)^k \geq 2.$$

Taking logarithms gives

$$k \log\left(\frac{1}{1-p}\right) \geq \log 2.$$

Notice $1/(1-p)$ is greater than 1, and so its logarithm is positive. Dividing out by that logarithm gives us an inequality on the k-values:

$$k \geq \frac{\log 2}{\log(1/(1-p))}.$$

Remark. The formula $P(X \geq k) = (1-p)^{k-1}$ was on the "discrete densities" handout. In this problem we have used it with k replaced by $k+1$.

5. [25=12+13 points] In a good winter there are 3 storms, on average. In a bad winter there are 6 storms, on average. Winters are good with probability 1/3 and bad with probability 2/3.

(a) Find a numerical formula for the probability that a 5-storm winter is bad (and explain what kind of random variables you are using).

Solution.

$$\begin{aligned} P(\text{bad}|5 \text{ storm}) &= \frac{P(\text{bad and } 5 \text{ storm})}{P(5 \text{ storm})} \\ &= \frac{P(5 \text{ storm}|\text{bad})P(\text{bad})}{P(5 \text{ storm}|\text{bad})P(\text{bad}) + P(5 \text{ storm}|\text{good})P(\text{good})} \\ &= \frac{e^{-6}\frac{6^5}{5!} \cdot (2/3)}{e^{-6}\frac{6^5}{5!} \cdot (2/3) + e^{-3}\frac{3^5}{5!} \cdot (1/3)}. \end{aligned}$$

Here storms occur randomly in time, and so we use Poisson random variables: a Poisson(3) random variable for the number of storms in a good winter, and a Poisson(6) random variable for the number of storms in a bad winter.

(b) Write X for the number of storms per winter. Then $E[X] = 3 \cdot (1/3) + 6 \cdot (2/3) = 5$. Show $E[X^2] = 32$, and then evaluate $\text{Var}(X)$.

Solution.

$$\begin{aligned} E[X^2] &= E[X^2|\text{bad}]P(\text{bad}) + E[X^2|\text{good}]P(\text{good}) \\ &= (6 + 6^2) \cdot (2/3) + (3 + 3^2) \cdot (1/3) \\ &= 32, \end{aligned}$$

where we have used that for a Poisson random variable $Y \sim \text{Poisson}(\lambda)$, one has $E[Y^2] = \text{Var}(Y) + (E[Y])^2 = \lambda + \lambda^2$.

Hence $\text{Var}(X) = E[X^2] - (E[X])^2 = 32 - 5^2 = 7$.

6. [25 points] An airplane has 85 seats for passengers. The airline has sold 100 tickets, but each passenger has only an 80% chance of showing up for the flight. (We assume passengers travel alone, and are independent of one another.)

Find the approximate probability that every passenger who shows up will get a seat on the airplane.

Solution. Write X for the number of passengers who show up, so that $X \sim \text{Binomial}(100, .80)$ because there are 100 passengers and we regard each one as a Bernoulli trial, with "success" meaning they show up for the flight. The mean and variance of X are

$$\mu = np = 100(.8) = 80, \qquad \sigma^2 = np(1-p) = 100(.8)(.2) = 16.$$

The Binomial density is difficult to work with (by hand) because the numbers are so large, and so instead we use the Normal approximation (which is valid since n is large and $\sigma^2 > 10$).

We want

$$\begin{aligned} P(X \le 85) &= P(X \le 85.5) \qquad \text{using the "continuity correction"} \\ &= P\left(\frac{X-\mu}{\sigma} \le \frac{85.5-\mu}{\sigma}\right) \\ &\simeq P\left(Z \le \frac{5.5}{\sqrt{16}}\right) \qquad \text{where } Z \text{ is standard normal} \\ &\simeq P(Z \le 1.375) \\ &\simeq .915 \end{aligned}$$

Incidentally, using the original Binomial random variable gives an answer (by computer) of about .920.

Remark. The Poisson approximation is not suitable for this problem, because $p = .8$ is large here.

7. [*Extra credit*, 20 points] An airplane has s seats for passengers. The airline has sold n tickets, but each passenger has only probability p of showing up for the flight. (We assume passengers travel alone, and are independent of one another.) Write r for the price of each ticket, c for the cost (to the airline) of each flight, and d for the cost (to the airline) of each passenger who shows up but cannot get a seat on the airplane. Let X denote the number of passengers who show up (a random variable). Then

$$\text{profit} = nr - c - d(X - s)_+$$

where $t_+ = t$ if $t > 0$ and $t_+ = 0$ if $t \leq 0$.

(a) Evaluate the expected profit, using the Normal approximation.

Solution. We have $X \sim \text{Binomial}(n, p)$ because there are n passengers and we regard each one as a Bernoulli trial, with "success" meaning they show up for the flight. The mean and variance of X are

$$\mu = np, \qquad \sigma^2 = np(1-p).$$

Therefore by linearity of expectation,

$$\begin{aligned} E[\text{profit}] &= nr - c - dE[(X-s)_+] \\ &= nr - c - d\sigma E\left[\left(\frac{X-\mu}{\sigma} - \frac{s-\mu}{\sigma}\right)_+\right] \\ &\simeq nr - c - d\sigma E\left[\left(Z - \frac{s-\mu}{\sigma}\right)_+\right] \\ &\qquad \text{by the Normal approximation to the Binomial} \\ &\simeq nr - c - d\sigma \int_{-\infty}^{\infty} \left(z - \frac{s-\mu}{\sigma}\right)_+ \phi(z)dz \end{aligned}$$

where $\phi(z)$ is the density function for the standard normal random variable. When evaluating the integral, notice we only really need to integrate from $z = \frac{s-\mu}{\sigma}$ to $z = \infty$.

Aside. We should really do a continuity correction in the above calculation. Where should it go?

(b) Explain in principle how the airline can determine the best number n of seats to sell.

Solution. The airline should try to choose n to maximize the expected profit. They could do this by evaluating the above formula for each n, or else by regarding n as a real variable (rather than an integer variable) and using calculus methods to find the maximum: differentiate the expected profit with respect to n and find the critical points, then check which critical point is the maximum, and then round n to the nearest integer.