Name: ______________________________

Signature: ______________________________

Math 5651. Lecture 001 (V. Reiner)   Final Exam
Friday, December 17, 2010

This is a 120 minute exam. No books, notes, calculators, cell phones or other electronic devices are allowed. There are a total of 100 points. To get full credit for a problem you must show the details of your work. Answers unsupported by an argument will get little credit. Do all of your calculations on this test paper.

Problem	Score
1.	______
2.	______
3.	______
4.	______
5.	______
6.	______
Total:	______

Problem 1. *(15 points total; 5 points each)* Recall that a Poisson random variable X with mean λ is one that has probability function $f(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, and that its variance is also λ.

Assume that traffic accidents occur at a certain intersection according to a Poisson *process* that averages one accident every 10 days.

a. (5 points) What is the probability of *no* accidents in a given *year*?

b. (5 points) What is the probability of *at least two* accidents in a given year. Express your answer without any summation symbols.

c. (5 points) What is the standard deviation in the number of accidents that occur in a given year?

Problem 2. *(15 points total)*

a. (8 points) Let X be the sum of 100 rolls of a fair 6-sided die having $1, 2, 3, 4, 5, 6$ on its sides. Compute the central limit theorem's approximation to the probability that this sum is at least 400. Express your answer in terms of the (cumulative) distribution function $\Phi(x)$ for a standard normal random variable.
(Hint: $1+2+3+4+5+6 = 21$ and $1^2+2^2+3^2+4^2+5^2+6^2 = 91$.)

b. (7 points) If X_1, X_2, X_3 are independent and identically distributed normal random variables of mean 10 and standard deviation 1, what is the probability that $2X_1 + 4X_2 > 5X_3$? Again, express your answer in terms of $\Phi(x)$.

Problem 3. *(15 points total; 5 points each)* Recall that an exponentially distributed random variable X with parameter β has probability density function $f(x) = \beta e^{-\beta x}$ for $x > 0$ and 0 for $x \leq 0$.

a. (5 points) Compute the median for such a random variable, as a function of the parameter β.

b. (5 points) Assume $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables having such an exponential distribution with parameter β, and let $Y = \max\{X_1, \ldots, X_n\}$. Compute the cumulative distribution function $F(y) = \Pr(Y \leq y)$ explicitly, as a function of y, n, β.

c. (5 points) Compute the probability density function $f(y)$ for Y explicitly, as a function of y, n, β.

Problem 4. *(15 points)* Recall that a random variable X having a gamma distribution with parameters α, β has moment generating function $\psi_X(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}$.

a. (7 points) Compute the third moment $\mu_3(X) = E(X^3)$ for such a random variable, as a function of α, β.

b. (8 points) Recall that the $\alpha = 1$ special case of the gamma distribution is the exponential distribution with parameter β.

Prove that if $X_1, X_2, \ldots, X_n$ are independent and identically distributed with exponential distribution of parameter β, then their sample mean $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ has a gamma distribution with some parameters $\overline{\alpha}, \overline{\beta}$. Say exactly what these parameters $\overline{\alpha}, \overline{\beta}$ are explicitly in terms of n and β.

Problem 5. *(15 points)* Prove the following statement, called the *Robbins Lemma*, for a Poisson[1] random variable X of mean λ: any random variable of the form $f(X)$ for a function f will satisfy

$$E\left(X \cdot f(X-1)\right) = \lambda \cdot E\, f(X).$$

(**NOTE**: During the exam, there was a typo that omitted the factor of X on the left side, making the statement entirely wrong, and so I did not grade this problem. Unfortunately, during the exam when people were pointing out that something had to be wrong, I didn't notice what the real problem was, and I tried to fix it incorrectly. It was only after the exam that we discovered the real problem.)

[1]... whose definition was recalled in Problem 1.

Problem 6. *(20 points total)* Recall that a random variable X having a beta distribution with parameters α, β has probability density function $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ for $x \in (0,1)$, and 0 for $x \notin (0,1)$

Your friend pulls a coin from her pocket, and claims that it was produced in a factory where the heads probability $P = p$ of the coins produced follows a beta distribution with parameters $\alpha = \beta = 2$. She tells you that she intends to flip the coin 100 times and count the number of heads as a random variable X.

a. (4 points) What is the marginal p.d.f. $f_1(p)$ for P?

b. (4 points) What is the conditional p.d.f $g_2(x|p)$ for X given $P = p$?

c. (4 points) What is the joint p.d.f. $f(p, x)$ for (P, X)?

d. (8 points) She then flips the coin 100 times, and heads appears 20 times total. Show that the conditional density $g_1(p|X = 20)$ for the heads probability P given that $X = 20$ again has a beta distribution for some parameters $\hat{\alpha}, \hat{\beta}$, and explicitly identify these parameters.

Brief solutions

1.The number of accidents X per year should be Poisson with mean $\lambda = \frac{1}{10} \cdot 365 = 36.5$. Hence one has...

(a) $\Pr(X = 0) = e^{36.5} \frac{36.5^0}{0!} = e^{36.5}$.

(b)

$$\begin{aligned} \Pr(X \geq 2) &= 1 - \Pr(X = 0) - \Pr(X = 1) \\ &= 1 - e^{-36.5} \frac{36.5^0}{0!} - e^{-36.5} \frac{36.5^1}{1!} \\ &= 1 - e^{-36.5} 37.5 \end{aligned}$$

(c) X has variance also $\lambda = 36.5$, so its standard deviation is $\sqrt{36.5}$.

2.(a) $X = X_1 + \cdots + X_{100}$ where the X_i are i.i.d. with

$$\begin{aligned} EX_i &= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2} \\ E(X_i^2) &= \frac{1}{6}\left(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2\right) = \frac{91}{6} \\ \mathrm{Var}(X_i) &= E(X_i^2) - (EX_i)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} \end{aligned}$$

Hence

$$EX = 100 \cdot \frac{7}{2} = 350$$

and

$$\mathrm{Var}(X) = 100 \cdot \frac{35}{12}, \qquad \sigma(X) = 10\sqrt{\frac{35}{12}} = 5\sqrt{\frac{35}{3}}.$$

The central limit theorem says X is approximately normally distributed, with the above mean and standard deviation, so

$$\begin{aligned}
\Pr(X \geq 400) &= \Pr\left(X - 350 \geq 50\right) \\
&= \Pr\left(\frac{X - 350}{5\sqrt{\frac{35}{3}}} \geq \frac{50}{5\sqrt{\frac{35}{3}}}\right) \\
&\approx \Pr\left(Z \geq \frac{50}{5\sqrt{\frac{35}{3}}}\right) \quad \text{for a standard normal } Z \\
&= 1 - \Pr\left(Z \leq \frac{10}{\sqrt{\frac{35}{3}}}\right) \\
&= 1 - \Phi\left(\frac{10}{\sqrt{\frac{35}{3}}}\right)
\end{aligned}$$

(b) $\Pr(2X_1 + 4X_2 > 5X_3) = \Pr(2X_1 + 4X_2 - 5X_3 > 0)$. So define

$$Y = 2X_1 + 4X_2 - 5X_3$$

which is normal with

$$EY = 2 \cdot 10 + 4 \cdot 10 - 5 \cdot 10 = 10$$

$$\operatorname{Var}(Y) = 2^2 \cdot 1 + 4^2 \cdot 1 + (-5)^2 \cdot 1 = 45, \qquad \sigma(Y) = \sqrt{45} = 3\sqrt{5}.$$

Hence

$$\begin{aligned}
\Pr(Y \geq 0) &= \Pr\left(Y - 10 \geq -10\right) \\
&= \Pr\left(\frac{Y - 10}{3\sqrt{5}} \geq \frac{-10}{3\sqrt{5}}\right) \\
&= \Pr\left(Z \geq \frac{-10}{3\sqrt{5}}\right) \quad \text{for a standard normal } Z \\
&= 1 - \Pr\left(Z \leq \frac{-10}{3\sqrt{5}}\right) \\
&= 1 - \Phi\left(\frac{-10}{3\sqrt{5}}\right) \left(= \Phi\left(\frac{10}{3\sqrt{5}}\right)\right)
\end{aligned}$$

3. (a) We need to solve for m in

$$\frac{1}{2} = \int_{-\infty}^{m} \beta e^{-\beta x} dx = \left[e^{-\beta x}\right]_{-\infty}^{m} = 1 - e^{-\beta m}$$

$$\text{so one has} \quad \frac{1}{2} = e^{-\beta m}$$

$$\log\left(\frac{1}{2}\right) = -\beta m$$

$$m = \frac{\log\left(\frac{1}{2}\right)}{-\beta} = \frac{\log 2}{\beta}$$

(b)

$$\begin{aligned} F(y) &= \Pr(Y \le y) \\ &= \Pr(X_1 \le y \text{ and } \cdots \text{ and } X_n \le y) \\ &= \Pr(X_1 \le y) \cdots \Pr(X_n \le y) \end{aligned}$$

and

$$\Pr(X_i \le y) = \int_{-\infty}^{y} \beta e^{-\beta x} dx = 1 - e^{-\beta y}$$

for $y \ge 0$ and 0 otherwise, so

$$F(y) = \left(1 - e^{-\beta y}\right)^n$$

for $y \ge 0$ and 0 otherwise.
(c) Y has pdf

$$f(y) = \frac{d}{dy} F(y) = n\left(1 - e^{-\beta y}\right)^n \left(\beta e^{-\beta y}\right)$$

4.

$$\begin{aligned} \psi_X(t) &= \left(\frac{\beta}{\beta - t}\right)^{\alpha} = \beta^{\alpha} \cdot (\beta - t)^{-\alpha} \\ \psi'_X(t) &= \alpha\beta^{\alpha} \cdot (\beta - t)^{-\alpha-1} \\ \psi''_X(t) &= \alpha(\alpha+1)\beta^{\alpha} \cdot (\beta - t)^{-\alpha-2} \\ \psi'''_X(t) &= \alpha(\alpha+1)(\alpha+2)\beta^{\alpha} \cdot (\beta - t)^{-\alpha-3} \\ \text{so that } \mu_3(X) &= E(X^3) = \psi'''_X(t = 0) \\ &= \alpha(\alpha+1)(\alpha+2)\beta^{\alpha}\,(\beta)^{-\alpha-3} \\ &= \frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} \end{aligned}$$

(b)

$$\begin{aligned}\Psi_{\overline{X}_n}(t) &= \Psi_{\frac{1}{n}X_1+\cdots+X_n}(t)\\ &= \Psi_{X_1+\cdots+X_n}\left(\frac{t}{n}\right)\\ &= \Psi_{X_1}\left(\frac{t}{n}\right)\cdots\Psi_{X_n}\left(\frac{t}{n}\right)\\ &= \left(\frac{\beta}{\beta-\frac{t}{n}}\right)^n\\ &= \left(\frac{n\beta}{n\beta-t}\right)^n\end{aligned}$$

which is the moment generating function for a gamma distribution with parameters $\overline{\alpha} = n$ and $\overline{\beta} = n\beta$.

5. This problem was not graded, but here is the solution to the corrected statement:

$$\begin{aligned}E\left(Xf(X-1)\right) &= \sum_{k=0}^{\infty} kf(k-1)e^{-\lambda}\frac{\lambda^k}{k!}\\ &= \sum_{k=1}^{\infty} kf(k-1)e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{since the } k=0 \text{ term vanishes}\\ &= \sum_{k=1}^{\infty} f(k-1)e^{-\lambda}\frac{\lambda\cdot\lambda^{k-1}}{(k-1)!}\\ &= \lambda\sum_{\ell=0}^{\infty} f(\ell)e^{-\lambda}\frac{\lambda^\ell}{\ell!} \quad \text{reindexing } \ell := k-1\\ &= \lambda Ef(X)\end{aligned}$$

6. (a)

$$f_1(p) = \frac{\Gamma(2+2)}{\Gamma(2)\Gamma(2)}p^{2-1}(1-p)^{2-1}f_1(p) = \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}p(1-p)$$

for p in $(0,1)$, and 0 otherwise.

(b)

$$g_2(x|p) = \binom{100}{x}p^x(1-p)^{100-x}$$

for $x = 0, 1, 2, \ldots, 100$.
(c)

$$\begin{aligned} f(p,x) &= g_2(x|p)f_1(p) \\ &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}p(1-p)\cdot\binom{100}{x}p^x(1-p)^{100-x} \\ &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}\binom{100}{x}p^{x+1}(1-p)^{101-x} \end{aligned}$$

for p in $(0, 1)$ and $x = 0, 1, 2, \ldots, 100$, and 0 otherwise.
(d)

$$\begin{aligned} g_1(p|x=20) &= \frac{f(p,x)}{f_2(20)} \\ &= \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)f_2(20)}\binom{100}{20}p^{21}(1-p)^{81} \end{aligned}$$

for p in $(0, 1)$. Since this conditional density is *proportional* to a beta distribution having parameters $\hat{\alpha} = 22$ and $\hat{\beta} = 82$, it must actually *equal* such a distribution, that is, the constant in front must normalize it properly.