Name:________________________ **ID:**

18440: Probability and Random variables
Quiz 2
Friday, November 14th, 2014

- You will have 50 minutes to complete this test.
- No calculators, notes, or books are permitted.
- If a question calls for a numerical answer, you do not need to multiply everything out. (For example, it is fine to write something like (0.9)7!/(3!2!) as your answer.)
- Don't forget to write your name on the top of every page.
- Please show your work and explain your answer. We will not award full credit for the correct numerical answer without proper explanation. Good luck!

Name:________________________ **ID:**

Problem 1 (25 points) Let $X_1, \ldots, X_n$ be independent standard normal variables $N(0,1)$.

- (5 points) What is the law of $\sum_{i=1}^n X_i$? (give the name and density)
- (5 points) What is the law of $\sum_{i=1}^n X_i^2$? (give the name and density)
- (5 points) Let (X,Y) be jointly Gaussian, that is with density

$$f(x,y) = Ce^{-\frac{x^2}{2}-\frac{y^2}{2}-cxy}$$

with C so that $\int f(x,y)dxdy = 1$. with $|c| < 1$. Compute $Cov(X,Y)$ and show that X and Y are independent iff $Cov(X,Y) = 0$.
- (10 points) Determine the joint density of $U = X_2$ and $V = X_1/X_2$ and show that V has a Cauchy law.

Answer:

- The sum of independent normal random variables is a normal random variable with mean and variance equal to the sum of the means and sum of the variances respectively of the component variables. So, $X = \sum_{i=1}^n X_i$ is a $N(0,n)$ and therefore has density

$$f_X(y) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{y^2}{2n}}$$

- $Y = \sum_{i=1}^n X_i^2$ is a chi-squared distribution, that is a $\Gamma(n/2, 1/2)$ distribution, which has density

$$\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}$$

- Notice that $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ by symmetry. Do the change of variable $x \to z = x - cy$ to find that

$$\int xye^{-\frac{x^2}{2}-\frac{y^2}{2}-cxy} dxdy = \int (z+cy)ye^{-\frac{z^2}{2}-(1-c^2)\frac{y^2}{2}} dzdy = c\sqrt{2\pi} \int y^2 e^{-(1-c^2)\frac{y^2}{2}} dy$$

Similarly $\int e^{-\frac{x^2}{2}-\frac{y^2}{2}-cxy} dxdy = \sqrt{2\pi} \int y^2 e^{-(1-c^2)\frac{y^2}{2}} dy$ Hence

$$Cov(X,Y) = \mathbb{E}[XY] = \frac{c \int y^2 e^{-(1-c^2)\frac{y^2}{2}} dzdy}{\int e^{-(1-c^2)\frac{y^2}{2}} dy} = (1-c^2)^{-1}c\,.$$

If $Cov(X,Y) = 0$, $c = 0$ by the above and the density is a product of a function of x by a function of y, and therefore X and Y are independent. If X, Y are independent, and centered, the covariance vanishes. Hence $Cov(X,Y) = 0$ iff X, Y are independent.
- Either compute the Jacobian for the change of variables or for a given X_2, remark that the law of $V = g(X_1) = X_1/X_2$ has density

$$C|X_2|e^{-X_2^2 V^2/2}$$

by the change of variable formula (here C is the normalizing constant). As a consequence, the joint law of (U,V) is given by

$$f(v,u) = f_{V|U}(v|x) f_U(x) = C|x|e^{-\frac{x^2}{2}(v^2+1)}$$

Then, the density of V is given by

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} f_{V,U}(v,x)dx \\ &= C\int_{-\infty}^{\infty} |x|xe^{-\frac{x^2}{2}(v^2+1)} \\ &= C'(v^2+1)^{-1} \end{aligned}$$

by rescaling. Here C' is some constant. This is a Cauchy distribution. and $C = \frac{1}{\pi}$

Name:__________________________ **ID:**

Problem 2 (30 points) At a bus stop, the times at which bus 69, bus 12, bus 56, bus 49 arrive are independent and form Poisson point processes with rate $\lambda = 6$/hour, 4/hour, 10/hour and 2/hour respectively.

- (5 points) Write down the probability density function for the amount of time until the first bus 69 arrives.
- (5 points) Let T be the first time one of the buses $69, 12, 49$ or 56 arrive. Write down the probability density function for T and name the distribution.
- (5 points) Compute the probability that exactly 5 bus 56 pass during the first hour.
- (5 points) Bus 69 and 12 go to the train station. Knowing that non of these 2 buses passed between noon and 1 PM, what is the probability one of them will come before 1 : 30 PM ?
- (10 points) Compute the probability that exactly 5 bus 56 pass before the first bus 69.

Answer:

- The pdf for the time until the first occurence of a Poisson process is exponential: $p(t) = \lambda \exp -\lambda t = 6e^{-6t}$.
- T is the minimum of a set of exponentially distributed variables, so it has an exponential distribution with parameter $\lambda = 6 + 4 + 10 + 2 = 22$.
$$p(t) = 22e^{-22t}$$
- The number of buses that arrive in a time interval t follows a Poisson distribution with parameter $\lambda t = 10$, so
$$P(5 \text{ bus } 56) = \frac{10^5}{5!}e^{-10}$$
- Exponential distributions are memoryless; this is the same as the probability that one comes in the first half hour.
$$\begin{aligned} P(69 \cup 12) &= 1 - P(\text{no } 69 \cap \text{no } 12) \\ &= 1 - P(\text{no } 69)P(\text{no } 12) \\ &= 1 - e^{-\frac{6}{2}}e^{-\frac{4}{2}} \\ &= 1 - e^{-5} \end{aligned}$$
- Letting E be the event that exactly 5 bus 56 arrive before the first bus 69, we can find $P(E)$ by conditioning on the time t when the first bus 69 arrives:
$$\begin{aligned} P(E) &= \int_0^\infty P(E|t)p(t)dt \\ &= \int_0^\infty \left(\frac{(10t)^5}{5!}e^{-10t}\right)(6e^{-6t})dt \\ &= \frac{6(10^5)}{5!}\int_0^\infty t^5 e^{-16t}dt \\ &= \frac{6(10^5)}{5!16^6}\int_0^\infty u^5 e^{-u}du \end{aligned}$$
To evaluate the integral, integrate by parts five times: each integration by parts brings down the exponent of u by one, so altogether this adds a factor of 5!. Then
$$P(E) = \frac{3}{8}\left(\frac{5}{8}\right)^5 \int_0^\infty e^{-u}du = \frac{3 \times 5^5}{8^6}$$

Name:_________________________ **ID:**

Problem 3 (15 points) Let X_1, X_2, X_3 be independent uniform variables on $[0,1]$. Let $X = \min\{X_1, X_2, X_3\}$ and $Y = \max\{X_1, X_2, X_3\}$

- (5 points) Compute the density function of (X, Y).
- (5 points) Compute $Cov(X, Y)$.
- (5 points) Are X and Y independent ? Why ?

Answer:

-
- We compute the joint distribution of X, Y. Take $x \leq y$. Note that by translation and rescaling

$$\begin{aligned} P(\{Y \leq y\} \cap \{X \geq x\}) &= \int_{x \leq x_1, x_2, x_3 \leq y} dx_1 dx_2 dx_3 \\ &= 3! \int_{x \leq x_1 \leq x_2 \leq x_3 \leq y} dx_1 dx_2 dx_3 \\ &= 3! \int_{0 \leq x_1 \leq x_2 \leq x_3 \leq y-x} dx_1 dx_2 dx_3 \\ &= 3!(y-x)^3 \int_{0 \leq x_1 \leq x_2 \leq x_3 \leq 1} dx_1 dx_2 dx_3 \\ &= (y-x)^3 \end{aligned}$$

If $y < x$, $P(\{Y \leq y\} \cap \{X \geq x\}) = 0$. so that

$$f_{X,Y}(x,y) = -\partial_x \partial_y P(\{Y \leq y\} \cap \{X \geq x\}) = 6(y-x) 1_{x \leq y} \,.$$

- Y and X have expected values $\mathbb{E}[Y] = \int_0^1 3y^3 dy = \frac{3}{4}$ and $\mathbb{E}[X] = 1 - \mathbb{E}[Y] = \frac{1}{4}$ respectively.

$$\begin{aligned} \mathbb{E}[XY] &= \int 6xy(y-x) dx dy \\ &= \int (3x^2y^2 - 2x^3y) \Big|_{x=0}^{x=y} dy \\ &= (\frac{1}{5}y^5) \Big|_0^1 \\ &= \frac{1}{5} \end{aligned}$$

So $Cov(X,Y) = \frac{1}{5} - \frac{1}{4}\frac{3}{4} = \frac{1}{80}$.

- They are not independent. The covariance is not zero, and the conditional probability $p(Y = y | X = x)$ does not equal the probability $p(Y = y)$.

Name:________________________ **ID:**

Problem 4 (20 points) Toss a coin independently 10000 times where the probability of a head is p.

- (5 points) What is the probability that the first head appears at the 1000^{th} toss ? Give the exact formula, name the distribution, and then approximate it when $p = 0,001$.
- (5 points) What is the probability that the tenth head appears at the 1000^{th} toss ? Give the exact formula, name the distribution, and then approximate it when $p = 0,001$.
- (10 points) Let Y be the number of heads till time 10000. Approximate $P(Y \geq 7689)$ when $p = 1/2$. You may use the function $\Phi(a) = \int_{-\infty}^{a} e^{-x^2/2}dx/\sqrt{2\pi}$.

Answer: Let h_n be the the toss when the n^{th} head appears.

- This is a negative binomial distribution:
$$p(h_1) = (1-p)^{999}p.$$
For $p = 0.001$, $p(h_1) = \frac{p}{1-p}(1-p)^{\frac{1}{p}} \approx \frac{0.001}{0.999}e^{-1} = \frac{1}{999e} \approx 0.00037$.
- This is also a negative binomial:
$$p(h_{10}) = \binom{999}{9}(1-p)^{990}p^{10}.$$
We can approximate it with a gamma distribution with parameters $t = 10$ and $\beta = p$, evaluated at $x = 1000$:
$$p(h_{10}) \approx \frac{\beta^t x^{t-1} e^{-\beta x}}{\Gamma(t)} = \frac{e^{-1}}{1000\Gamma(10)} = \frac{1}{1000(9!)e}$$
- The number of heads has a binomial distribution. When $p = 1/2$, this is approximately a normal distribution with mean $\mu = np = 5000$ and variance $\sigma^2 = np(1-p) = 2500$. So
$$P(Y \geq 7689) \approx \int_{\mu + \frac{2688.5}{50}\sigma}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{2688.5}{50}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Phi(\frac{2688.5}{50}).$$
This is very close to zero.

Name:________________________ **ID:**

Problem 5 (10 points) Let W be a Gamma random variable with parameters (t, β) and suppose that conditional on $W = w$, $X_1, \ldots, X_n$ are independent exponential variables with rate w. Show that the conditional distribution of W given $X_1 = x_1, X_2 = x_2 \ldots, X_n = x_n$ is gamma with parameter $(t + n, \beta + \sum_{i=1}^n x_i)$

Answer:
We know $p(x_1 \ldots x_n|w)$ and $p(w)$, so to find $p(w|x_1 \ldots x_n)$ we just need to use Bayes' rule:

$$\begin{aligned}
p(w|x_1 \ldots x_n) &= \frac{p(x_1 \ldots x_n|w)p(w)}{p(x_1 \ldots x_n)} \\
&= w^n e^{-w \sum_{i=1}^n x_i} \frac{\beta^t w^{t-1} e^{-\beta w}}{\Gamma(t)} \left(\int_0^\infty (w')^n e^{-w' \sum_{i=1}^n x_i} \frac{\beta^t (w')^{t-1} e^{-\beta w}}{\Gamma(t)} dw' \right)^{-1} \\
&= w^{t+n-1} e^{-w(\beta + \sum_{i=1}^n x_i)} \left(\int_0^\infty (w')^{n+t-1} e^{-w'(\beta + \sum_{i=1}^n x_i)} dw' \right)^{-1} \\
&= w^{t+n-1} e^{-w(\beta + \sum_{i=1}^n x_i)} \left((\beta + \sum_{i=1}^n x_i)^{-n-t} \int_0^\infty (u')^{n+t-1} e^{-u} du \right)^{-1} \\
&= \frac{w^{t+n-1} (\beta + \sum_{i=1}^n x_i)^{n+t} e^{-w(\beta + \sum_{i=1}^n x_i)}}{\Gamma(t+n)}.
\end{aligned}$$

This is a gamma distribution with the given parameters.