STATISTICS 251
HOMEWORK ASSIGNMENT 7
DUE FRIDAY NOVEMBER 10

Problem 1. Discretization of an Exponential. Let T have an exponential distribution with parameter $\lambda > 0$. Define Y to be the largest integer smaller than T. What is the distribution of Y, that is, for any integer $m \geq 0$, what is $P\{Y = m\}$?

Problem 2. Independent Gaussians. Suppose that X and Y are independent random variables each with the standard normal distribution. Let $X = R\cos\Theta$ and $Y = R\sin\Theta$ be the polar coordinate representation of the point (X, Y), with the angular coordinate Θ chosen so that $0 \leq \Theta < 2\pi$.

(A) Find the density of Y/X. HINT: This might be related to something on HW 6.
(B) Show that $\tilde{X} = R\cos 2\Theta$ and $\tilde{Y} = R\sin 2\Theta$ are independent standard normal random variables.
(C) Use (B) to show that the random variables

$$\frac{2XY}{\sqrt{X^2+Y^2}} \quad \text{and} \quad \frac{X^2-Y^2}{\sqrt{X^2+Y^2}}$$

are independent standard normal random variables. NOTE: You can either look up the relevant trig double-angle formulas or you can forget all the trig you ever knew and just learn Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Problem 3. Addition of Gammas. The Gamma density with shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$ is

$$\begin{aligned} f_{\alpha,\lambda}(x) &= \tfrac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{for } x \geq 0, \\ &= 0 \quad \text{for } x < 0. \end{aligned}$$

The constant $\Gamma(\alpha)$ is a normalizing constant. (When α is a positive integer, $\Gamma(\alpha) = (\alpha - 1)!$.)

(A) Show that if X and Y are independent random variables with densities $f_{\alpha,\lambda}$ and $f_{\beta,\lambda}$, respectively, then $X + Y$ has density $f_{\alpha+\beta,\lambda}$. NOTE: We did the case where α, β are integers in class.
(B) Show that if X is standard normal then X^2 has density $f_{1/2,1/2}$.
(C) Show that if $X_1, X_2, \ldots, X_n$ are independent standard normal then $\sum_{i=1}^n X_i^2$ has density $f_{n/2,1/2}$.

Problem 4. Chi-square distribution. For $m = 1, 2, \ldots$ let R_{2m}^2 be a random variable with the chi-square distribution on $2m$ degrees of freedom. Use the connection between the gamma distribution and the Poisson process to find formulas (in terms of Poisson probabilities) for

(A) the c.d.f. of R_{2m}^2; and
(B) the c.d.f. of R_{2m}.

Problem 5. A box contains n balls labeled $1, 2, 3, \ldots, n$. Balls are drawn one at a time at random, with replacement after each draw, until the first draw that produces a ball obtained on some previous draw. Let D_n be the (random) number of draws required. (Thus, D_n can take any integer value between 2 and $n + 1$.)

(A) Show that for each fixed value of $x > 0$,

$$\lim_{n \to \infty} P\{D_n/\sqrt{n} > x\} = e^{-x^2/2}.$$

(B) The continuous probability distribution with c.d.f.

$$F(x) = 1 - e^{-x^2/2} \quad \text{for } x \geq 0,$$
$$= 0 \quad \text{for } x \leq 0$$

is called the *Rayleigh* distribution. Part (A) shows that for large n the random variable $D_n/\sqrt{n}$ is approximately distributed as a Rayleigh random variable. Calculate the expectation of a Rayleigh random variable.

Problem 6. The sunrise problem. Quality control at the U. S. Mint (where new coins are manufactured) has fallen to such an astoundingly low level that the $p-$value of a new coin (that is, the probability that when tossed the particular coin will land H) could be anything between 0 and 1. Assume, then, that the $p-$value of a new coin is a random variable Θ that is uniformly distributed on the unit interval. When this coin comes into your possession you have no idea what its $p-$value Θ is. So like any good statistician, you toss the coin n times in succession and obtain S_n Heads.

(A) What is the probability that $S_n = k$, for $k = 0, 1, 2, \dots, n$?
(B) Give an explanation of your answer to (A) that involves no integrals and no need to look up the Binomial distribution. HINT: Find a way to "simulate" your coin-tossing experiment using $n + 1$ independent Uniform- (0,1) random variables.
(C) Conditional on the event that $S_n = k$, what is the probability that $\Theta \leq x$? NOTE: Your answer can be left as the ratio of two integrals.
(D) What is the name of the distribution you got in part (C)? HINT: Take d/dx to get a density, then find this density somewhere in Pitman, maybe sec. 4.6.

Historical Note: Why is it called the "sunrise problem"? The answer has something to do with the origins of what is now called "Bayesian" statistics. In the mid-1700s, Pierre-Simon Laplace used the problem as an illustration of how one might make inference from experimental data. He asked the question: "Given that I have witnessed the sun rise on n successive days, what is the probability that it will rise tomorrow"? He then suggested that on any given day the event of a successful sunrise should be treated as a coin toss with a $p-$coin whose $p-$value is a random variable Θ with the uniform- (0,1) distribution. (This is a bit of an over-simplification: Laplace realized that it was not realistic to view the events of sunrise on successive days as conditionally independent.)