Review Problems for Exam 1

1. Let X and Y be independent normal$(0,1)$ random variables and define $W = \dfrac{(X-Y)^2}{2}$. Give the distribution of W.

   Suggestion: First, determine the distribution of $X - Y$.

2. Let X denote a random variable with mgf $M_X(t)$ defined on some interval around 0. Put $S(t) = \ln(M_X(t))$ and prove that
$$S'(0) = E(X) \quad \text{and} \quad S''(0) = \text{var}(X).$$

3. A median of a distribution of a random variable, X, is a value, m, such that
$$\text{P}(X \leqslant m) \geqslant \frac{1}{2} \quad \text{and} \quad \text{P}(X \geqslant m) \geqslant \frac{1}{2}.$$

   (a) Prove that if X is continuous with pdf f_X, then a median m satisfies
$$\int_{-\infty}^{m} f_X(x)\, \mathrm{d}x = \int_{m}^{+\infty} f_X(x)\, \mathrm{d}x = \frac{1}{2}.$$

   (b) Let $\beta > 0$ and $X \sim$ exponential(β). Compute a median of X. Is the value you obtained the only median of the distribution? How does your answer compare with the mean of the distribution?

   (c) Show that if X is a continuous random variable, and m is a median of the the distribution of X, then m a number which minimizes the expression
$$h(t) = E(|X - t|) \quad \text{for} \ \ t \in \mathbb{R}.$$
   That is, $E(|X - m|) = \min\limits_{t \in \mathbb{R}} E(|X - t|)$.

4. Give a random variable, X, of expected value μ and variance σ^2, the *skewness* of the distribution of X, denoted Skew(X), is defined to be
$$\text{Skew}(X) = \frac{E(X-\mu)^3}{\sigma^3}.$$

   (a) Let $\beta > 0$ and $X \sim$ exponential(β). Compute a skewness of X.

   (b) Let $Z \sim$ normal$(0,1)$. Compute the skewness of Z.

5. Let X and Y be independent, normal$(0, \sigma^2)$ random variables, and define

$$U = X^2 + Y^2 \quad \text{and} \quad V = \frac{X}{\sqrt{U}}.$$

   (a) Find the joint pdf, $f_{(U,V)}$, of U and V.

   (b) Show that U and V are independent random variables.

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pdf f_X, and let $\overline{X}_n$ denote the sample mean. Prove that the pdf of the sample mean satisfies

$$f_{\overline{X}_n}(t) = n f_Y(nt), \quad \text{for all } \ t \in \mathbb{R},$$

   where $Y = \displaystyle\sum_{i=1}^{n} X_i$.

7. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Gamma$(2, \theta)$ distribution, where θ is an unknown parameter. Define $Y = \displaystyle\sum_{i=1}^{n} X_i$.

   (a) Find the distribution of Y and determine c so that the statistic $T = cY$ is an unbiased estimator for θ.

   (b) If $n = 5$, show that

$$\mathrm{P}\left(9.59 < \frac{2Y}{\theta} < 34.2\right) \approx 0.95.$$

   (c) Use Part (b) to show that if a sample of size $n = 5$ is collected from a Gamma$(2, \theta)$ distribution, and the sum of the values of the sample is y, then the interval

$$\left(\frac{2y}{34.2}, \frac{2y}{9.59}\right)$$

   is a 95% confidence interval for θ.

   (d) Suppose the values in a random sample of size $n = 5$ from a Gamma$(2, \theta)$ distribution are:

   44.8079 1.5215 12.1929 12.5734 43.2305

   Use the data to obtain a point estimate for θ and a 95% confidence interval for θ.

   Give an interpretation of your result.

8. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal(μ, σ^2) distribution and define the statistic

$$T_n = \sum_{i=1}^{n} (X_i - \overline{X}_n)^2,$$

where $\overline{X}_n$ denotes the sample mean. We will show later in this course that $\dfrac{1}{\sigma^2} T_n$ has a χ^2 distribution with $n - 1$ degrees of freedom.

   (a) Explain how you would use knowledge of the distribution of $\dfrac{1}{\sigma^2} T_n$ to obtain a $100(1-\alpha)\%$ confidence interval for the variance σ^2 of a normal(μ, σ^2) distribution based on a random sample of size n from that distribution.

   (b) Give a 90% confidence interval for the variance of a normal(μ, σ^2) distribution based on the statistic T_n, where the sample size, n, is 20.

9. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with unknown expectation, μ, and unknown variance, σ^2. Define the statistic

$$T_n = \sum_{i=1}^{n} (X_i - \overline{X}_n)^2,$$

where $\overline{X}_n$ denotes the sample mean.

   (a) Starting with

$$(X_i - \mu)^2 = [(X_i - \overline{X}_n) + (\overline{X}_n - \mu)]^2,$$

   where $\overline{X}_n$ denotes the sample mean, derive the identity

$$\sum_{i=1}^{n} (X_i - \mu)^2 = T_n + n(\overline{X}_n - \mu)^2. \tag{1}$$

   (b) Take expectations on both sides of equation (1) to derive a formula for $E(T_n)$ in terms of σ^2. Is T_n an unbiased estimator for σ^2?