Ground rules: You must work alone on this quiz. Do not collaborate with anyone, either within or outside the class, to obtain answers or even hints. Questions of clarification should be directed to the instructor; in addition, the use of the Internet should be avoided. This quiz contains 6 questions, each of equal weight, making the quiz worth 60 points.

1. Suppose Y is a continuous random variable with probability density function (pdf)

$$f_Y(y) = \begin{cases} k(4y - 2y^2), & 0 < y < 2 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the value of k and then graph the pdf.
(b) Find the cumulative distribution function (cdf) of Y and graph it. Remember to consider the different cases (similarly to how we did it in the notes/HW).
(c) Compute $P(0.2 < Y < 0.7)$ and $P(Y = 1)$.
Note: I want very detailed graphs.

2. The **median** of a continuous random variable Y with cdf $F_Y(y)$ is the value m such that $F_Y(m) = P(Y \leq m) = 0.5$. That is, Y is just as likely to be larger than its median as it is to be smaller. If $f_Y(y)$ denotes the pdf of Y, then we know that m solves the following equation

$$0.5 = F_Y(m) = \int_{-\infty}^{m} f_Y(y)dy;$$

that is, the area under $f_Y(y)$ to the left of m is 0.5 and the area to the right of m is also 0.5. For each of the distributions, calculate the median m.
(a) $Y \sim \mathcal{U}(0, \theta)$
(b) $Y \sim \text{exponential}(\beta)$
(c) $Y \sim \mathcal{N}(\mu, \sigma^2)$.
Note: For each of these distributions, graph the pdf and mark the median on the horizontal axis.

3. Suppose that Y is a continuous random variable with pdf

$$f_Y(y) = \begin{cases} 2y/\theta^2, & 0 < y < \theta \\ 0, & \text{otherwise.} \end{cases}$$

In this pdf, θ is a positive constant.
(a) Find $E(Y)$ and $V(Y)$.
(b) Find the median of Y.
Note: In both parts, your answers should depend on θ.

4. Human papillomavirus (HPV) infection has been established as the cause of virtually all forms of cervical cancer. Large cohort studies incorporate active follow-up with multiple visits and the collection of cervical specimens for HPV DNA testing. This is an important part of disease assessment and allows researchers to collect important covariate

information that is linked to disease status. In a recent study, the Guanacaste Project, one of covariates recorded (for all eligible subjects) was Y, the age at which the subject gave birth to her first child. This variable is possibly linked to HPV infectivity because HPV is a sexually transmitted disease. A normal probability model is attached to Y with mean $\mu = 27.3$ years and standard deviation $\sigma = 3.5$ years. Use this probability model to answer the following questions.
(a) Find the probability that a mother gives birth to her first child after she is 35 years old.
(b) Find the proportion of first-time mothers between 20 and 30 years.
(c) Ten percent of first-time mothers give birth before what age?
Note: In each part, draw detailed picture.

5. Suppose that the random variable Y has the following pdf

$$f_Y(y) = \begin{cases} ke^{-y^2/2}, & 0 < y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

Note that Y does not have a normal distribution because the support $R = \{y : y > 0\}$. In fact, Y is said to have a **folded-normal** distribution.
(a) Argue that $k = 2/\sqrt{2\pi}$. *Hint:* For this part, use the fact that the standard normal density is a symmetric function and integrates to 1.
(b) Show that $E(Y) = 2/\sqrt{2\pi}$ and $V(Y) = 1 - 2/\pi$.

6. Let Y be a random variable with mean $\mu = E(Y)$ and variance $\sigma^2 = V(Y)$. Recall from Quiz 2 that the **skewness** associated with Y, denoted by ξ, is given by

$$\xi = \frac{E[(Y-\mu)^3]}{\sigma^3}.$$

The skewness ξ quantifies the level of which a probability distribution departs from symmetry. Another measure that describes the distribution of Y is the kurtosis. The **kurtosis** associated with Y, denoted by κ, is given by

$$\kappa = \frac{E[(Y-\mu)^4]}{\sigma^4}.$$

The kurtosis κ measures the "peakedness" of a distribution; i.e., how tall the probability density function is at its peak. A normal random variable Y, for example, has $\kappa = 3$ irrespective of its mean or standard deviation. If a random variable's kurtosis is greater than 3, it is said to be **leptokurtic**. If its kurtosis is less than 3, it is said to be **platykurtic**. Leptokurtosis is associated with pdfs that are simultaneously "peaked" and have "fat tails." Platykurtosis is associated with pdfs that are simultaneously less peaked and have thinner tails.
(a) Derive the skewness and kurtosis associated with $Y \sim \text{exponential}(\beta)$. Follow the hint given in Quiz 2 for computing the skewness. The kurtosis is computed in a similar

way (with the appropriate modifications). Is this distribution skewed right or skewed left? Is this distribution leptokurtic or platykurtic? Use your numerical values of ξ and κ to answer this question.