4.6 The Gamma Distribution and Related Distributions

- Many variables have distributions that are nonnegative and skewed to the right (long right tail).

Examples:
- Lifelengths of manufactured parts
- Lengths of time between arrivals at a restaurant
- Survival times for severely ill patients

- Such r.v.'s may often be modeled with a gamma distribution.
- We will study the gamma distribution in general and two special cases of it.

Defn: A continuous r.v. Y has a gamma distribution [shorthand: $Y \sim \text{Gamma}(\alpha, \beta)$] if its pdf is

where $\alpha > 0$, $\beta > 0$ and

$\Gamma(\alpha) =$

is called the gamma function evaluated at α.

Graphs of specific gamma pdf's:

Gamma(1,1) pdf:

Gamma(2,1) pdf:

- We see the shape of the pdf changes for different values of α.
- Hence α is called the ________ parameter of the gamma density.
- Also, β is called the ________ parameter of the gamma density.

Facts about the gamma function

(i)

(ii)

Example:

(iii)

Example: What is $\Gamma(\frac{5}{2})$?

Theorem: The gamma(α, β) pdf is a valid density.

Proof:

Note: The kernel of a pdf $f(y)$ is the part that depends on y.

- We see that the gamma pdf (like any pdf) consists of a kernel and a "normalizing constant."
- This constant is free of y, but it forces the pdf to integrate to 1 over the support.

Gamma pdf:

- This fact will be quite helpful to us.

<u>Theorem</u> (Gamma Mean and Variance): If $Y \sim \text{gamma}(\alpha, \beta)$, then

<u>Proof</u>:

Theorem: (Gamma mgf): If $Y \sim \text{Gamma}(\alpha, \beta)$, then the mgf of Y is

Proof:

Example 1: Suppose Y is a r.v. with pdf

$$f(y) = \begin{cases} c\,y^4 e^{-y/3} & \text{if } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

for some constant c.

- What value of c makes $f(y)$ a valid density?

- What is $E(Y)$? What is $V(Y)$?

Gamma Probabilities

- For any $a < b$ where $a, b > 0$, the probability

  $P(a < Y < b) =$

cannot be found via direct integration unless α is an integer.

- Such gamma probabilities can be found in R (see examples on course web page).

Example 1 again: Suppose $Y \sim \text{Gamma}(5, 3)$. Find the probability that Y falls within 2 standard deviations of its mean.

- Two special cases of the gamma distribution are especially important in statistics.

Defn. For any integer $\nu \geq 1$, a r.v. Y has a chi-square (χ^2) distribution with ν degrees of freedom [Shorthand: $Y \sim \chi^2(\nu)$] if Y is a gamma r.v. with $\alpha = \frac{\nu}{2}$ and $\beta = 2$.

- The χ^2 distribution occurs often in theoretical statistics.
- Values of $\chi^2(\nu)$ quantiles are given in Table 6 of Appendix 3.

Theorem: If $Y \sim \chi^2(\nu)$, then

Proof: This follows immediately from the formulas for the gamma mean and variance.

The Exponential Distribution

- The exponential distribution is commonly used as a model for lifetimes, survival times, or waiting times.

Defn. A r.v. Y has an exponential distribution [Shorthand: $Y \sim \text{expon}(\beta)$] if Y is a gamma r.v. with $\alpha = 1$.

- The exponential pdf is (for $\beta > 0$):

- The exponential cdf can be found analytically:

Plot of expon ($\beta = 2$) pdf:

Plot of expon ($\beta = 2$) cdf:

Theorem: If $Y \sim \text{expon}(\beta)$, then

Theorem: If $Y \sim \text{expon}(\beta)$, then the mgf of Y is:

Proofs: Let $\alpha = 1$ in the gamma formulas.

Example 2: Let the lifetime of a part (in thousands of hours) follow an exponential distribution with mean lifetime 2000 hours ($\beta =$).

- Find the probability the part lasts less than 1500 hours.

- Find the probability the part lasts more than 2200 hours.

- Find the probability the part lasts between 2000 and 2500 hours.

Memoryless Property of Exponential

- Let a lifelength Y follow an exponential distribution. Suppose the part has lasted "a" units of time already. Then the (conditional) probability of it lasting at least b additional units of time is

the same as the probability of lasting at least b units of time in the first place.

Proof:

- The exponential distribution is the only continuous distribution with this "memoryless" property.
- However, the geometric distribution also has this memoryless property.

Relationship with Poisson process

- Suppose we have a Poisson process with a mean of λ events per time unit. Let $\beta = \frac{1}{\lambda}$. Consider the waiting time W until the first occurrence. Then $W \sim \text{expon}(\beta)$.

Proof: Clearly W is continuous with support on $[0, \infty)$. For any $w \geq 0$, note that if the number of occurrences in $[0,1]$ is Poisson with mean λ, then the number of occurrences in $[0, w]$ is

- Note: This result can be generalized: For any integer $\alpha \geq 1$, the waiting time W until the α-th occurrence has a

Example 3: Suppose customers arrive in a queue according to a Poisson process with mean $\lambda = 12$ per hour. What is the probability we must wait more than 10 minutes until the first customer?

- What is the probability we must wait more than 10 minutes to see the fourth customer?