Exponential Random Variables: The Poisson Connection

Suppose that calls are received at a 24-hour "suicide hotline" according to a Poisson process with rate α = .5 call per day.

Then the number of days X between successive calls has an exponential distribution with parameter value .5, so the probability that more than 2 days elapse between calls is

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) \\ &= 1 - F(2; .5) \\ &= e^{-(.5)(2)} \\ &= .368 \end{aligned}$$

The expected time between successive calls is 1/.5 = 2 days.

161

The Gamma Function

To define the family of gamma distributions, we first need to introduce a function that plays an important role in many branches of mathematics.

Definition: For $\alpha > 0$, the **gamma function** $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

162

The Gamma Function

The most important properties of the gamma function are the following:

1. For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$ [via integration by parts]

2. For any positive integer, n, $\Gamma(n) = (n - 1)!$

3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

163

The Gamma Function

If we let $$f(x; \alpha) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

then f(x; α) ≥ 0 and $\int_0^\infty f(x; \alpha)\, dx = \Gamma(\alpha)/\Gamma(\alpha) = 1$

so $f(x$; a) satisfies the two basic properties of a pdf.

164

The Gamma Distribution

Definition: A continuous random variable X is said to have a **gamma distribution** if the pdf of X is

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where the parameters α and β satisfy $\alpha > 0$, $\beta > 0$.

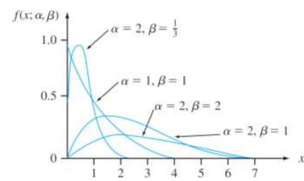
The **standard gamma distribution** has $\beta = 1$, so the pdf of a standard gamma rv is given by

$$f(x; \alpha) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

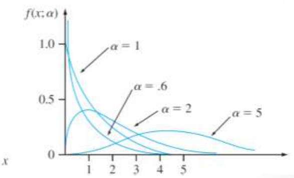
165

The Gamma Distribution

The exponential distribution results from taking $\alpha = 1$ and $\beta = 1/\lambda$. The figure on the left illustrates the graphs of the gamma pdf $f(x; \alpha, \beta)$ for several (α, β) pairs, whereas the figure on the right presents graphs of the standard gamma pdf.



Gamma density curves



standard gamma density curves

166

The Gamma Distribution

For the standard pdf, when $\alpha \leq 1$, $f(x; \alpha)$, is strictly decreasing as x increases from 0; when $\alpha > 1$, $f(x; \alpha)$ rises from 0 at $x = 0$ to a maximum and then decreases.

The parameter β is called the *scale parameter* because values other than 1 either stretch or compress the pdf in the x direction.

167

The Gamma Distribution

The mean and variance of a random variable X having the gamma distribution $f(x; \alpha, \beta)$ are

$$E(X) = \mu = \alpha\beta \qquad V(X) = \sigma^2 = \alpha\beta^2$$

When X is a standard gamma rv, the cdf of X,

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}\, dy \qquad x > 0$$

is called the **incomplete gamma function.**

168

The Gamma Distribution

The incomplete gamma function can also be used to compute probabilities involving nonstandard gamma distributions. These probabilities can also be obtained almost instantaneously from various software packages.

Proposition

Let *X* have a gamma distribution with parameters α and β. Then for any $x > 0$, the cdf of *X* is given by

$$P(X \leq x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$$

where $F(\bullet\,; \alpha)$ is the incomplete gamma function.

169

The Gamma Distribution

Probabilities for the Gamma distribution are best computed using software. We will use R.

The CDF of a Gamma is given using the pgamma function in R.

So, F(x, α, β) = pgamma(x, α, 1/β).

R uses 1/β as the rate, λ, of the underlying Poisson process.

We can use the standard gamma:

pgamma(x, α, 1/β) = pgamma(x /β, α)

170

Example

Suppose the reaction time *X* of a randomly selected individual to a certain stimulus has a standard gamma distribution ($\beta = 1$) with $\alpha = 2$.

Since $P(a \leq X \leq b) = F(b) - F(a)$

we have $P(3 \leq X \leq 5) = F(5; 2,1) - F(3; 2,1)$

=pgamma(5,2) – pgamma(3,2)

= 0.9595723 – 0.8008517

= 0.1587206

171

Example

The probability that the reaction time is more than 4 sec is

$P(X > 4) = 1 - P(X \leq 4)$

$= 1 - F(4; 2)$

= 1 – pgamma(4,2)

= 1 – 0.9084218

= 0.0915782

172

Example

Suppose the survival time X in weeks of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 15$.

(Data in *Survival Distributions: Reliability Applications in the Biomedical Services,* by A. J. Gross and V. Clark, suggests $\alpha \approx 8.5$ and $\beta \approx 13.3$.)

The expected survival time is $E(X) = (8)(15) = 120$ weeks, whereas $V(X) = (8)(15)^2 = 1800$ and $\sigma_x = \sqrt{1800} = 42.43$ weeks.

173

Example

The probability that a mouse survives between 60 and 120 weeks is

$$P(60 \le X \le 120) = P(X \le 120) - P(X \le 60)$$
$$= F(120/15;\, 8) - F(60/15;\, 8)$$
$$= F(8;8) - F(4;8)$$
$$= \text{pgamma}(8,8) - \text{pgamma}(4,8)$$
$$= 0.5470392 - 0.05113362$$
$$= 0.4959056$$

174

Example

The probability that a mouse survives at least 30 weeks is

$$P(X \ge 30) = 1 - P(X < 30)$$
$$= 1 - P(X \le 30)$$
$$= 1 - F(30/15;\, 8)$$
$$= 1 - \text{pgamma}(2,8)$$
$$= 1 - 0.001096719$$
$$= 0.9989033$$

175

The Chi-Squared Distribution

The chi-squared distribution is important because it is the basis for a number of procedures in statistical inference.

The central role played by the chi-squared distribution in inference springs from its relationship to normal distributions.

We will see the chi-squared distribution later in the course.

176

The Chi-Squared Distribution

Definition: Let v be a positive integer. Then a random variable X is said to have a **chi-squared distribution** with parameter v if the pdf of X is the gamma density with $\alpha = v/2$ and $\beta = 2$. The pdf of a chi-squared rv is thus

$$f(x;\nu) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The parameter is called the **number of degrees of freedom** (df) of X. The symbol χ^2 is often used in place of "chi-squared."

177

The beta distribution

The **beta** distribution models random variables that take the value [0,1]

It arises naturally as the proportional ratio of two gamma distributed random variables

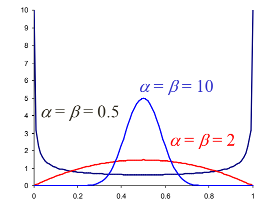
$$X \sim \Gamma(\alpha_1, \theta)$$

$$Y \sim \Gamma(\alpha_2, \theta)$$

$$\frac{X}{X+Y} \sim \text{beta}(\alpha_1, \alpha_2)$$

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

The expectation is $\alpha/(\alpha + \beta)$



178