XXI. Gamma Distribution (with Exponential and Chi-square)

Gamma Function

- The gamma function is

$$\boxed{\Gamma(\alpha) := \int_0^\infty y^{\alpha-1} e^{-y}\, dy}.$$

- Properties:
  - $\boxed{\Gamma(n+1) = n\Gamma(n)}$
  - When n is a whole number, it satistfies

$$\boxed{\Gamma(n) = (n-1)!}.$$

(Plug in a value for α, and it spits out a number.)

Don't mix up the gamma function and the gamma distribution!

Formula for the Gamma Distribution

- Fixed parameters: $\alpha > 0, \beta > 0$
- Density function:

$$\boxed{f(y) := \frac{y^{\alpha-1} e^{-\frac{y}{\beta}}}{\beta^\alpha \Gamma(\alpha)}, 0 \le y < \infty}$$

Don't mix up the gamma function and the gamma distribution!

Key Properties of the Gamma Distribution

- **Mean**:

$$\boxed{\mu = E(Y) = \alpha\beta}$$

- **Variance**:

$$\boxed{\sigma^2 = V(Y) = \alpha\beta^2}$$

- **Standard deviation**:

$$\boxed{\sigma = \sqrt{V(Y)} = \sqrt{\alpha\beta^2}}$$

Exponential Distribution

- The exponential distribution measures the waiting time until a random event.
- It is a special case of the gamma distribution with $\alpha := 1$ and β arbitrary.
- **Density**: $\boxed{f(y) = \frac{1}{\beta}e^{-\frac{y}{\beta}}, 0 \le y < \infty}$
- **Mean**: $\boxed{\mu = E(Y) = \beta}$
- **Variance**: $\boxed{\sigma^2 = V(Y) = \beta^2}$
- **Standard deviation**: $\boxed{\sigma = \beta}$

Chi-square Distribution

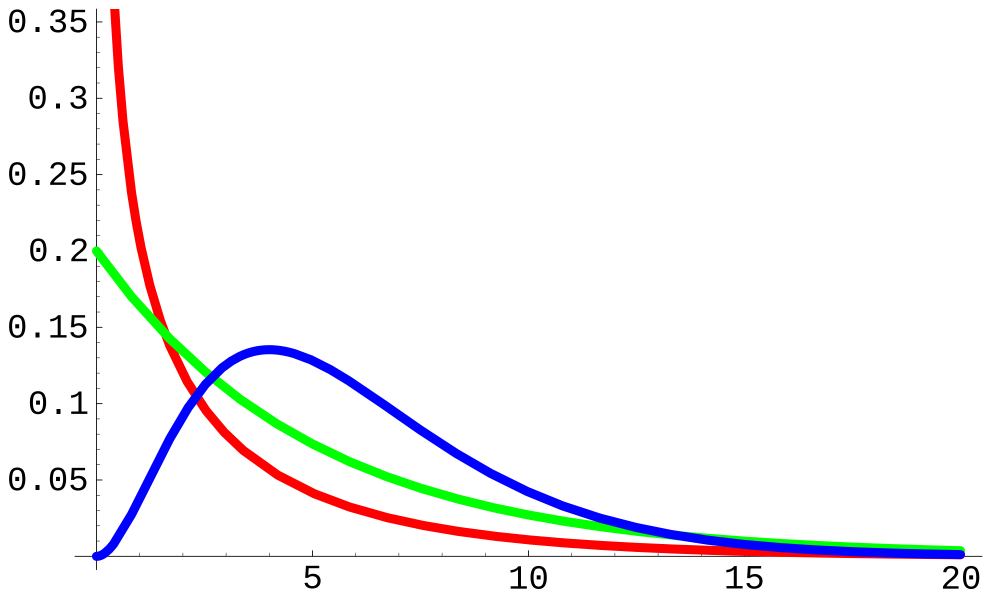
- The chi-square (χ^2) distribution is very useful in statistics.
- Let ν be a whole number. We have a chi-square distribution with ν degrees of freedom.
- It is a special case of the gamma distribution with $\alpha := \frac{\nu}{2}$ and $\beta := 2$.
- **Mean**: $\boxed{\mu = E(Y) = \nu}$
- **Variance**: $\boxed{\sigma^2 = V(Y) = 2\nu}$
- **Standard deviation**: $\boxed{\sigma = \sqrt{2\nu}}$

Example I

On the same set of axes, graph the density functions for the gamma distribution for the following combinations of (α, β):

$$\left(\frac{1}{2}, 5\right) \qquad (1, 5) \qquad (3, 2)$$

Describe the effects of changing α and β on the shape of the graph.



[$(\alpha, \beta) := \left(\frac{1}{2}, 5\right)$ goes to ∞ at 0. $(\alpha, \beta) := (1, 5)$ is exponential and goes to 1 at 0. $(\alpha, \beta) = (3, 2)$ goes to 0 at 0.]
α controls the <u>shape</u>. β just stretches it out to the right (and necessarily down, because the total area must be 1). Think of α as a <u>qualitative</u> difference and β as just a <u>quantitative</u> scalar.

Example II

Actuaries calculate that a car parked on the streets of Long Beach will stolen every twelve years on average. Use the exponential distribution to find the chance that your car will last 24 years without being stolen.

$Y :=$ time to be stolen
$f(y) = \frac{1}{\beta} e^{-\frac{y}{\beta}}, 0 \leq y < \infty$
Mean is $\beta = 12$.

$$\begin{aligned} P(Y > c) &= \int_c^\infty \frac{1}{\beta} e^{-\frac{y}{\beta}}\, dy \\ &= -e^{-\frac{y}{\beta}} \Big|_{y=c}^{y=\infty} \\ &= e^{-\frac{c}{\beta}} \\ &= e^{-\frac{24}{12}} = e^{-2} \approx 0.135 = \boxed{13.5\%} \end{aligned}$$

Example III

Seismic data indicate that the time until the next major earthquake in California is exponentially distributed with a mean of 10 years. Estimate the chance that there will be an earthquake in the next 30 years:

A. Use Markov's Inequality.

B. Use Tchebysheff's Inequality.

A. Markov says that $P(Y \geq 30) \leq \frac{E(Y)}{30} = \frac{1}{3}$, so the probability that there <u>will</u> be one is $\boxed{\geq \frac{2}{3}}$.

B. Since this is an exponential distribution, $\mu = \sigma = 10$. Tchebysheff tells us that

$$P(|Y - \mu| \geq 2\sigma) \leq \frac{1}{4},$$

so the probability that there <u>will</u> be one is $\boxed{\geq \frac{3}{4}}$.

Example IV

Seismic data indicate that the time until the next major earthquake in California is exponentially distributed with a mean of 10 years. Find the exact probability that there will be an earthquake in the next 30 years.

$$\int_0^{30} \frac{1}{10} e^{-\frac{y}{10}}\, dy = -e^{-\frac{y}{10}}\Big|_{y=0}^{y=30} = 1 - e^{-3} \approx 0.9502 \approx \boxed{95\%}$$

Example V

Let Y be an exponential distribution and let d and

m be constants. Prove that

$$P(Y > d + m | Y > d) = P(Y > m)$$

and interpret why the exponential distribution is called "memoryless".

$$\begin{aligned} P(Y > d + m | Y > d) &= \frac{e^{-\frac{(d+m)}{\beta}}}{e^{-\frac{d}{\beta}}} \\ &= e^{-\frac{m}{\beta}} \\ &= P(Y > m) \end{aligned}$$

Interpretation: If your car isn't stolen today, you get a fresh start tomorrow: Your probability tomorrow of surviving a month later (given that you make it through today) is the same as your probability today of surviving another month.
[The probability doesn't build up. If you get lucky today, you're just lucky – it doesn't mean you're "due" for a theft. So maybe this isn't a perfect match for earthquakes, because tectonic pressure does actually build up.]