18.05 Lecture 22
April 4, 2005

Central Limit Theorem
$X_1, ..., X_n$ - independent, identically distributed (i.i.d.)
$\overline{x} = \frac{1}{n}(X_1 + ... + X_n)$
$\mu = \mathbb{E}X, \sigma^2 = \text{Var}(X)$

$$\frac{\sqrt{n}(\overline{x} - \mu)}{\sigma} \xrightarrow[n \to \infty]{} N(0,1)$$

You can use the knowledge of the standard normal distribution to describe your data:

$$\frac{\sqrt{n}(\overline{x} - \mu)}{\sigma} = Y, \overline{x} - \mu = \frac{\sigma Y}{\sqrt{n}}$$

This expands the law of large numbers:
It tells you exactly how much the average value and expected vales should differ.

$$\frac{\sqrt{n}(\overline{x} - \mu)}{\sigma} = \sqrt{n}\frac{1}{n}(\frac{x_1 - \mu}{\sigma} + ... + \frac{x_n - \mu}{\sigma}) = \frac{1}{\sqrt{n}}(Z_1 + ... + Z_n)$$

where: $Z_i = \frac{X_i - \mu}{\sigma}; \mathbb{E}(Z_i) = 0, \text{Var}(Z_i) = 1$
Consider the m.g.f., see that it is very similar to the standard normal distribution:

$$\mathbb{E}e^{t\frac{1}{\sqrt{n}}(Z_1 + ... + Z_n)} = \mathbb{E}e^{tZ_1/\sqrt{n}} \times ... \times e^{tZ_n/\sqrt{n}} = (\mathbb{E}e^{tZ_1/\sqrt{n}})^n$$

$$\mathbb{E}e^{tZ_1} = 1 + t\mathbb{E}Z_1 + \frac{1}{2}t^2\mathbb{E}Z_1^2 + \frac{1}{6}t^3\mathbb{E}Z_1^3 + ...$$

$$= 1 + \frac{1}{2}t^2 + \frac{1}{6}t^3\mathbb{E}Z_1^3 + ...$$

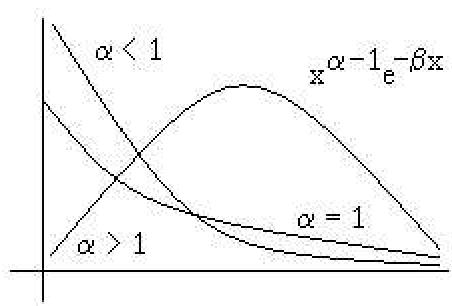
$$\mathbb{E}e^{(t/\sqrt{n})Z_1} = 1 + \frac{t^2}{2n} + \frac{t^3}{6n^{3/2}}\mathbb{E}Z_1^3 + ... \approx 1 + \frac{t^2}{2n}$$

Therefore:

$$(\mathbb{E}e^{tZ_1/\sqrt{n}})^n \approx (1 + \frac{t^2}{2n})^n$$

$$(1 + \frac{t^2}{2n})^n \xrightarrow[n \to \infty]{} e^{t^2/2} \text{ - m.g.f. of standard normal distribution!}$$

Gamma Distribution:
Gamma function; for $\alpha > 0, \beta > 0$



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

p.d.f of Gamma distribution, f(x):

$$1 = \int_0^\infty \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} dx, f(x) = \{\frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x}, x \geq 0; 0, x < 0\}$$

Change of variable $x = \beta y$, to stretch the function:

$$1 = \int_0^\infty \frac{1}{\Gamma(\alpha)} \beta^{\alpha-1} y^{\alpha-1} e^{-\beta y} \beta dy = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} dy$$

p.d.f. of Gamma distribution, $f(x|\alpha, \beta)$:

$$f(x|\alpha, \beta) = \{\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, x \geq 0; 0, x < 0\} - Gamma(\alpha, \beta)$$

Properties of the Gamma Function:

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

Integrate by parts:

$$= x^{\alpha-1} e^{-x}|_0^\infty - \int_0^\infty (-e^{-x})(\alpha-1) x^{\alpha-2} dx = 0 + (\alpha-1) \int_0^\infty x^{\alpha-2} e^{-x} dx = (\alpha-1)\Gamma(\alpha-1)$$

In summary, Property 1: $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

You can expand Property 1 as follows:

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = (n-1)(n-2)(n-3)\Gamma(n-3) =$$

$$= (n-1)...(1)\Gamma(1) = (n-1)!\Gamma(1), \Gamma(1) = \int_0^\infty e^{-x} dx = 1 \to \Gamma(n) = (n-1)!$$

In summary, Property 2: $\Gamma(n) = (n-1)!$

Moments of the Gamma Distribution:
$X \sim (\alpha, \beta)$

$$\mathbb{E}X^k = \int_0^\infty x^k \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{(\alpha+k)-1} e^{-\beta x} dx$$

Make this integral into a density to simplify:

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}} \int_0^\infty \frac{\beta^{\alpha+k}}{\Gamma(\alpha+k)} x^{(\alpha+k)-1} e^{-\beta x} dx$$

The integral is just the Gamma distribution with parameters $(\alpha+k, \beta)$!

$$= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\beta^k} = \frac{(\alpha+k-1)(\alpha+k-2) \times ... \times \alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta^k} = \frac{(\alpha+k-1) \times ... \times \alpha}{\beta^k}$$

For k = 1:

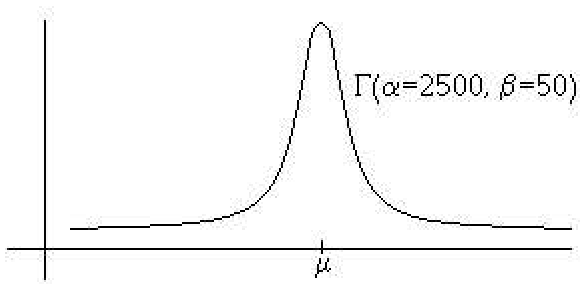
$$\mathbb{E}(X) = \frac{\alpha}{\beta}$$

For k = 2:

$$\mathbb{E}(X^2) = \frac{(\alpha+1)\alpha}{\beta^2}$$

$$\text{Var}(x) = \frac{(\alpha+1)\alpha}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}$$

Example:



If the mean = 50 and variance = 1 are given for a Gamma distribution,
Solve for $\alpha = 2500$ and $\beta = 50$ to characterize the distribution.

Beta Distribution:

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, 1 = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}dx$$

Beta distribution p.d.f. - $f(x|\alpha,\beta)$

Proof:

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty x^{\alpha-1}e^{-x}dx \int_0^\infty y^{\beta-1}e^{-y}dy = \int_0^\infty \int_0^\infty x^{\alpha-1}y^{\beta-1}e^{-(x+y)}dxdy$$

Set up for change of variables:

$$x^{\alpha-1}y^{\beta-1}e^{-(x+y)} = x^{\alpha-1}((x+y)-x)^{\beta-1}e^{-(x+y)} = x^{\alpha-1}(x+y)^{\beta-1}(1-\frac{x}{x+y})^{\beta-1}e^{-(x+y)}$$

Change of Variables:

$$s = x+y, t = \frac{x}{x+y}, x = st, y = s(1-t) \rightarrow Jacobian = s(1-t) - (-st) = s$$

Substitute:

$$= \int_0^1 \int_0^\infty t^{\alpha-1}s^{\alpha+\beta-2}(1-t)^{\beta-1}e^{-s}sdsdt = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt \int_0^\infty s^{\alpha+\beta-1}e^{-s}ds =$$

$$= \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \times \Gamma(\alpha+\beta) = \Gamma(\alpha)\Gamma(\beta)$$

Moments of Beta Distribution:

$$\mathbb{E}X^k = \int_0^1 x^k \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+k)-1}(1-x)^{\beta-1}dx$$

Once again, the integral is the density function for a beta distribution.

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+\beta+k)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} = \frac{(\alpha+k-1)\times ... \times \alpha}{(\alpha+\beta+k-1)\times ... \times (\alpha+\beta)}$$

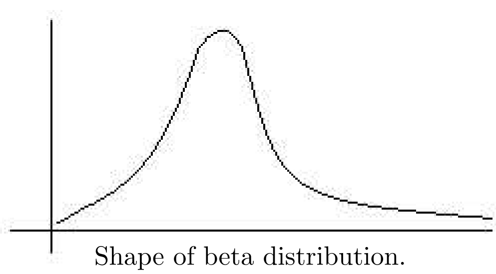
For k = 1:

$$\mathbb{E}X = \frac{\alpha}{\alpha+\beta}$$

For k = 2:

$$\mathbb{E}X^2 = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}$$

$$\text{Var}(X) = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$



Shape of beta distribution.

** End of Lecture 22