Chapter 16

Chi-Square, Weibull, Lognormal, Beta

Lecture 16 | Devore 9e: 4.4–4.5

Last time: change of variable, exponential, gamma. Today we finish the gamma and collect the remaining named continuous families. Every one of them is built out of pieces we already have.

16.1 Gamma, continued

Devore writes the gamma density in the equivalent form

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta\,\Gamma(\alpha)}\left(\dfrac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}, & x \geq 0, \\ 0, & x < 0, \end{cases} \qquad \mathbb{E}[X] = \alpha\beta, \quad \mathrm{Var}(X) = \alpha\beta^2.$$

Proposition 16.1 (Scaling)**.** *If* $X \sim \mathrm{Gamma}(\alpha,\beta)$ *then* $\dfrac{X}{\beta} \sim \mathrm{Gamma}(\alpha,1)$.

This is the change of variable of Lecture 15 with $g(x) = x/\beta$: the inverse is $g^{-1}(u) = \beta u$ with $\left|\frac{d}{du}g^{-1}(u)\right| = \beta$, so

$$f_{X/\beta}(u) = \beta f_X(\beta u) = \beta \cdot \frac{(\beta u)^{\alpha-1}e^{-u}}{\beta^\alpha \Gamma(\alpha)} = \frac{u^{\alpha-1}e^{-u}}{\Gamma(\alpha)}.$$

So, just as with the normal, one table suffices.

Definition 16.2 (Incomplete gamma function)**.** For the *standard* gamma ($\beta = 1$),

$$\mathrm{IGF}(x;\alpha) := \mathbb{P}\{X \leq x\} = \frac{1}{\Gamma(\alpha)}\int_0^x y^{\alpha-1}e^{-y}\,\mathrm{d}y.$$

Then for general β, by the scaling property,

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

which is looked up in a table exactly as Φ was.

Theorem 16.3 (Additivity)**.** *If* $X_1 \sim \mathrm{Gamma}(\alpha_1,\beta)$ *and* $X_2 \sim \mathrm{Gamma}(\alpha_2,\beta)$ *are indepen-*

dent (same β!), then

$$X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta).$$

This is just the waiting-time reading: waiting for α_1 arrivals and then for α_2 more is waiting for $\alpha_1 + \alpha_2$ arrivals. It also explains what a *fractional* α means: any $\text{Gamma}(\alpha, \beta)$ splits as a sum of a $\text{Gamma}(\lfloor \alpha \rfloor, \beta)$ piece and a $\text{Gamma}(\alpha - \lfloor \alpha \rfloor, \beta)$ piece.

Remark 16.4 (Shape)**.** $0 < \alpha < 1$ gives a density with a spike at 0 — the "early failure" regime, appropriate for items that fail immediately or not at all. $\alpha > 1$ gives an interior mode.

16.2 The chi-square distribution

Definition 16.5 (Chi-square)**.** Let $Z_1, \ldots, Z_k$ be i.i.d. $\text{Normal}(0, 1)$ and set

$$Y = Z_1^2 + Z_2^2 + \cdots + Z_k^2.$$

Then $Y \sim \chi^2_k$, chi-square with k **degrees of freedom**, with pdf

$$f(y; k) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma(k/2)} y^{k/2-1} e^{-y/2}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

Comparing with the gamma density:

$$\boxed{\chi^2_k = \text{Gamma}\left(\frac{k}{2}, 2\right)} \qquad \Longrightarrow \qquad \mathbb{E}[Y] = \frac{k}{2} \cdot 2 = k, \qquad \text{Var}(Y) = \frac{k}{2} \cdot 2^2 = 2k.$$

(For $k = 1$ this matches the transformation $Y = Z^2$ we did in Lecture 15.)

Remark 16.6 (Intuition: energy)**.** A gas molecule has velocity $\vec{V} = (V_x, V_y, V_z)$ with each component approximately $\text{Normal}(0, 1)$ after scaling. Its kinetic energy is proportional to

$$V_x^2 + V_y^2 + V_z^2 \sim \chi^2_3.$$

Chi-square is what you get when you add up squared fluctuations — which is exactly what a sample variance does, and why this distribution runs the statistics half of the course.

16.3 The Weibull distribution

Definition 16.7 (Weibull)**.** $X \sim \text{Weibull}(\alpha, \beta)$ with shape $\alpha > 0$ and scale $\beta > 0$ if

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\alpha}{\beta}\left(\dfrac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^\alpha}, & x \ge 0, \\ 0, & x < 0. \end{cases}$$

Its cdf is unusually simple:

$$F(x) = 1 - e^{-(x/\beta)^\alpha}, \qquad x \ge 0.$$

Remark 16.8 (Weibull versus gamma — look at the exponent)**.**

$$\text{Weibull:} \quad \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\boxed{\alpha}}}$$

$$\text{Gamma:} \quad \frac{1}{\beta\Gamma(\alpha)}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\boxed{1}}}$$

The polynomial factor is the same and the constants out front differ only in what it takes to normalize; the families differ in whether α appears in the exponent of the exponential. That one change is what makes the Weibull cdf elementary while the gamma cdf requires a table.

> **Theorem 16.9** (Where it comes from)**.** *If* $T \sim \text{Exponential}(1)$ *and* $X := \beta T^{1/\alpha}$*, then* $X \sim \text{Weibull}(\alpha, \beta)$*.*

Proof. $\mathbb{P}\{X \le x\} = \mathbb{P}\{\beta T^{1/\alpha} \le x\} = \mathbb{P}\{T \le (x/\beta)^\alpha\} = 1 - e^{-(x/\beta)^\alpha}$. □

So the Weibull is a *stretched* exponential waiting time. Setting $\alpha = 1$ recovers $\text{Exponential}(1/\beta)$ exactly.

> **Theorem 16.10.**
>
> $$\mathbb{E}[X] = \beta\,\Gamma\left(1 + \frac{1}{\alpha}\right), \qquad \operatorname{Var}(X) = \beta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma\left(1+\frac{1}{\alpha}\right)^2\right].$$

Remark 16.11 (Why engineers use it)**.** The exponential is memoryless, so it cannot describe wear-out. The Weibull can, through α:

$\alpha < 1$	failure rate *decreasing* — infant mortality
$\alpha = 1$	constant failure rate — exponential, no aging
$\alpha > 1$	failure rate *increasing* — wear-out

This is the standard lifetime model in reliability engineering.

16.4 The lognormal distribution

> **Definition 16.12** (Lognormal)**.** $X \sim \text{Lognormal}(\mu, \sigma^2)$ if $\ln X \sim \text{Normal}(\mu, \sigma^2)$.

Derive the density by the usual cdf route: for $x > 0$,

$$\mathbb{P}\{X \le x\} = \mathbb{P}\{\ln X \le \ln x\} = \Phi\left(\frac{\ln x - \mu}{\sigma}\right),$$

and differentiating (chain rule, $\frac{d}{dx}\ln x = \frac{1}{x}$),

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \cdot \frac{1}{x}, \qquad x > 0.$$

Theorem 16.13.
$$\mathbb{E}[X] = e^{\mu+\sigma^2/2}, \qquad \mathrm{Var}(X) = \big(e^{\sigma^2} - 1\big)e^{2\mu+\sigma^2}.$$

Remark 16.14 (Warning)**.** μ and σ^2 are the mean and variance of $\ln X$, *not* of X. Note $\mathbb{E}[X] = e^{\mu+\sigma^2/2} > e^{\mu}$, which is the median. The lognormal is right-skewed.

Example 16.15 (Multiplicative growth)**.** A stock grows by a random percentage each day:
$$X = X_0(1 + R_1)(1 + R_2)\cdots(1 + R_n).$$
Take logs:
$$\ln X = \ln X_0 + \sum_{i=1}^{n} \ln(1 + R_i).$$
The right-hand side is a *sum* of many small independent terms, so by the central limit theorem it is approximately normal — and therefore X itself is approximately lognormal.

Remark 16.16. So: normal is what you get from adding many small effects; lognormal is what you get from *multiplying* them. Prices, incomes, particle sizes, and file sizes are lognormal for this reason.

16.5 The beta distribution

Everything so far has had unbounded support. For a random *proportion* we need a distribution living on $[0, 1]$.

Definition 16.17 (Standard beta)**.** $X \sim \mathrm{Beta}(\alpha, \beta)$ if
$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1 - x)^{\beta-1}, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

To place it on a general interval $[A, B]$: say Y is beta on $[A, B]$ if $\frac{Y-A}{B-A} \sim \mathrm{Beta}(\alpha, \beta)$.

Theorem 16.18.
$$\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta}, \qquad \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Theorem 16.19 (Where it comes from)**.** *Let* $Y_1 \sim \mathrm{Gamma}(\alpha, 1)$ *and* $Y_2 \sim \mathrm{Gamma}(\beta, 1)$ *be independent. Then*
$$X := \frac{Y_1}{Y_1 + Y_2} \sim \mathrm{Beta}(\alpha, \beta).$$

Remark 16.20. That is the right way to read the parameters: α and β are the "sizes" of two competing gamma quantities, and X is the *share* taken by the first. For example, if spring rainfall is $\mathrm{Gamma}(\alpha, 1)$ and the rest of the year is $\mathrm{Gamma}(\beta, 1)$, then the fraction of the annual total falling in spring is $\mathrm{Beta}(\alpha, \beta)$. Note $\mathrm{Beta}(1, 1) = \mathrm{Uniform}[0, 1]$.

16.6 Summary: the continuous families

Distribution	**Relation**	**What it models**
Uniform$[A, B]$	—	equally likely anywhere in an interval
Normal(μ, σ^2)	—	sums of many small independent effects
Exponential(λ)	$=$ Gamma$(1, 1/\lambda)$	waiting time for the 1st arrival; memoryless lifetime
Gamma(α, β)	—	waiting time for α arrivals
χ^2_k	$=$ Gamma$(k/2, 2)$	sum of k squared standard normals; "energy"
Weibull(α, β)	$\beta T^{1/\alpha}$, $T \sim$ Exponential(1)	stretched waiting time; lifetimes *with* aging
Lognormal(μ, σ^2)	$\ln X \sim$ Normal(μ, σ^2)	multiplicative growth
Beta(α, β)	$\frac{Y_1}{Y_1+Y_2}$, Y_i gamma	a random proportion in $[0, 1]$

Remark 16.21. Notice how few genuinely independent objects there are. Starting from a Poisson process you get exponential, gamma, chi-square, Weibull and beta; starting from the normal you get chi-square and lognormal. Learning the *relations* is far more economical than memorizing eight densities.