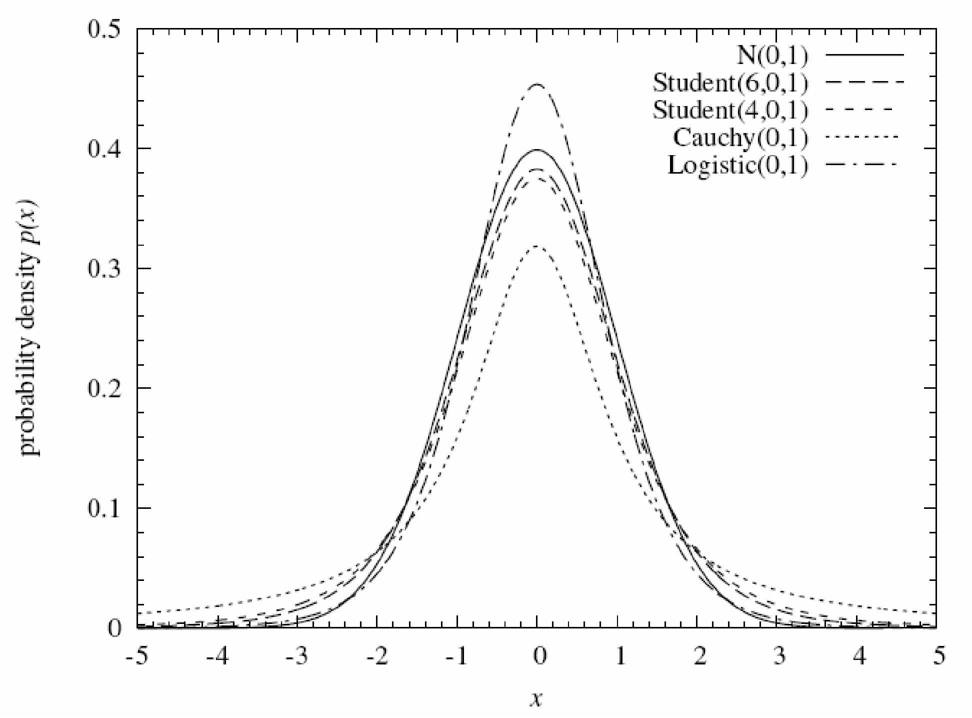
Opinionated Lessons in Statistics

by Bill Press

#8 Some Standard Distributions

Let us review some standard (i.e., frequently occurring) distributions:

The “bell shaped” ones differ qualitatively by their tail behaviors:



Normal (Gaussian) has the fastest falling tails:

$$x \sim \mathrm{N}(\mu, \sigma), \qquad \sigma > 0$$

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right)$$

Cauchy (aka Lorentzian) has the slowest falling tails:

$$x \sim \mathrm{Cauchy}(\mu, \sigma), \qquad \sigma > 0$$

$$p(x) = \frac{1}{\pi\sigma}\left(1 + \left[\frac{x-\mu}{\sigma}\right]^2\right)^{-1}$$

Cauchy has area=1 (zeroth moment), but no defined mean or variance (1st and 2nd moments divergent).

Student has power-law tails:

$$t \sim \text{Student}(\nu, \mu, \sigma), \qquad \nu > 0, \ \sigma > 0$$

$$p(t) = \frac{\Gamma(\frac{1}{2}[\nu+1])}{\Gamma(\frac{1}{2}\nu)\sqrt{\nu\pi}\sigma}\left(1 + \frac{1}{\nu}\left[\frac{t-\mu}{\sigma}\right]^2\right)^{-\frac{1}{2}(\nu+1)}$$

"bell shaped" but you get to specify the power with which the tails fall off. Normal and Cauchy are limiting cases. (Also occurs in some statistical tests.)

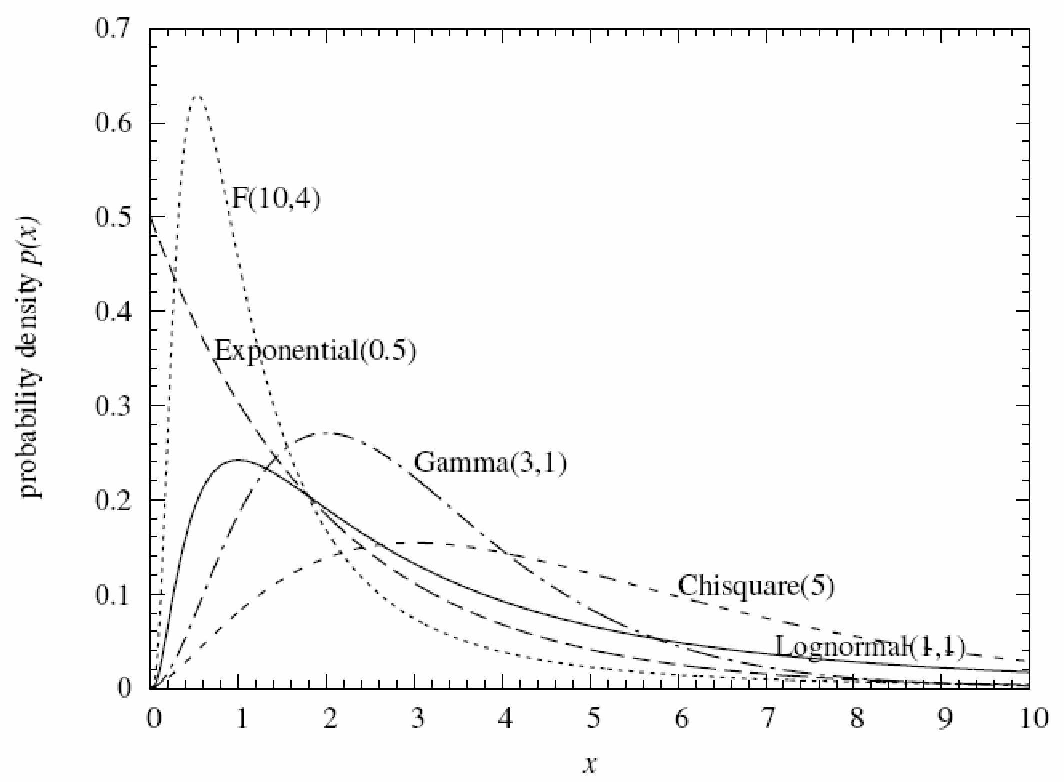
note that σ is not (quite) the standard deviation:

$$\text{Var}\{\text{Student}(\nu, \mu, \sigma)\} = \frac{\nu}{\nu - 2}\sigma^2$$

we'll see uses for "heavy-tailed" distributions later

"Student" was actually William Sealy Gosset (1876-1937), who spent his entire career at the Guinness brewery in Dublin, where he rose to become the company's Master Brewer. Brewing was one of the first "exact" modern manufacturing processes. More on Student later…

Another class of distributions model positive quantities:



Exponential:

$$x \sim \text{Exponential}(\beta), \qquad \beta > 0$$
$$p(x) = \beta \exp(-\beta x), \qquad x > 0$$

Lognormal:

$$x \sim \text{Lognormal}(\mu, \sigma), \qquad \sigma > 0$$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{1}{2}\left[\frac{\log(x) - \mu}{\sigma}\right]^2\right), \qquad x > 0 \tag{6.14.31}$$

Note the required extra factor of x^{-1} in front of the exponential: The density that is "normal" is $p(\log x)d \log x$.

While μ and σ are the mean and standard deviation in $\log x$ space, they are *not* so in x space. Rather,

$$\begin{aligned} \text{Mean}\{\text{Lognormal}(\mu, \sigma)\} &= e^{\mu + \frac{1}{2}\sigma^2} \\ \text{Var}\{\text{Lognormal}(\mu, \sigma)\} &= e^{2\mu}\, e^{\sigma^2}(e^{\sigma^2} - 1) \end{aligned} \tag{6.14.32}$$

```
p = (1 / Sqrt[2 Pi]) (1 / (sig x)) Exp[- (1 / 2) * (Log[x] - mu) ^2 / sig^2]
```

$$\frac{e^{-\frac{(-mu+Log[x])^2}{2\,sig^2}}}{\sqrt{2\pi}\ sig\ x}$$

```
moments = Integrate[{1, x, x^2} p, {x, 0, Infinity}, Assumptions → sig > 0,
  GenerateConditions → False]
```

$$\left\{1,\ e^{mu+\frac{sig^2}{2}},\ e^{2\,(mu+sig^2)}\right\}$$

```
Simplify[moments[[3]] - moments[[2]] ^2]
```

$$e^{2\,mu+sig^2}\left(-1 + e^{sig^2}\right)$$

Mathematica (and also MATLAB) can do these integrals, no problem!

Gamma distribution:

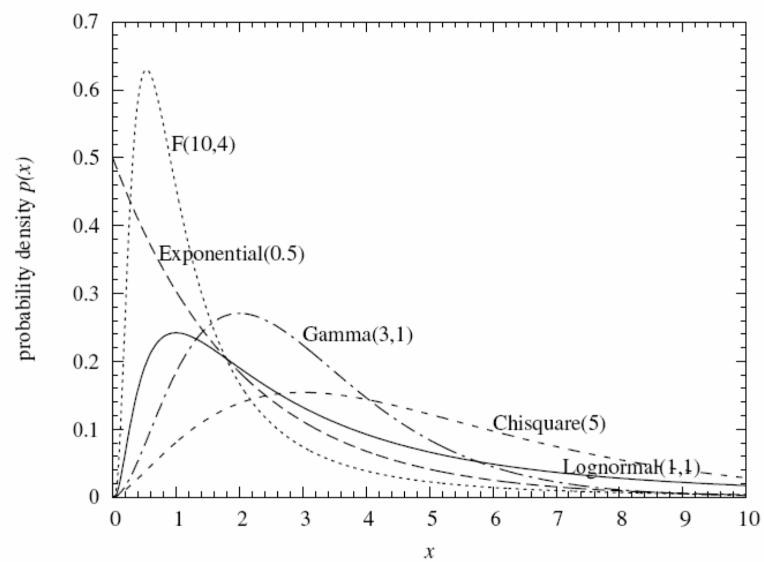
$$x \sim \text{Gamma}(\alpha, \beta), \qquad \alpha > 0,\ \beta > 0$$

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0$$

$$\text{Mean}\{\text{Gamma}(\alpha, \beta)\} = \alpha/\beta$$

$$\text{Var}\{\text{Gamma}(\alpha, \beta)\} = \alpha/\beta^2$$

When $\alpha \geq 1$ there is a single mode at $x = (\alpha - 1)/\beta$



- Gamma and Lognormal are both commonly used as convenient 2-parameter fitting functions for “peak with tail” positive distributions.
- Both have parameters for peak location and width.
- Neither has a separate parameter for how the tail decays.
  - Gamma: exponential decay
  - Lognormal: long-tailed (exponential of square of log)

Chi-square distribution (we'll use this a lot!)

Has only one parameter ν that determines both peak location and width.
ν is often an integer, called "number of degrees of freedom" or "DF"

the independent variable is χ^2, not χ

$$\chi^2 \sim \text{Chisquare}(\nu), \qquad \nu > 0$$

$$p(\chi^2)d\chi^2 = \frac{1}{2^{\frac{1}{2}\nu}\Gamma(\frac{1}{2}\nu)}(\chi^2)^{\frac{1}{2}\nu-1}\exp\left(-\tfrac{1}{2}\chi^2\right)d\chi^2, \qquad \chi^2 > 0$$

It's actually just a special case of Gamma, namely Gamma(ν/2,1/2)

$$\text{Mean}\{\text{Chisquare}(\nu)\} = \nu$$

$$\text{Var}\{\text{Chisquare}(\nu)\} = 2\nu$$

When $\nu \geq 2$ there is a single mode at $\chi^2 = \nu - 2$

Computationally, one wants efficient methods for all of:

- PDF *p(x)*
- CDF *P(x)*
- Inverse of CDF *x(P)*
- Random deviates drawn from it (we'll get to soon)

$$P(x) \equiv \int_{-\infty}^{x} p(x')dx'$$

NR3 has classes for many common distributions, with algorithms for p, cdf, and inverse cdf.

```
struct Normaldist : Erf {
Normal distribution, derived from the error function Erf.
    Doub mu, sig;
    Normaldist(Doub mmu = 0., Doub ssig = 1.) : mu(mmu), sig(ssig) {
    Constructor. Initialize with μ and σ. The default with no arguments is N(0, 1).
        if (sig <= 0.) throw("bad sig in Normaldist");
    }
    Doub p(Doub x) {
    Return probability density function.
        return (0.398942280401432678/sig)*exp(-0.5*SQR((x-mu)/sig));
    }
    Doub cdf(Doub x) {
    Return cumulative distribution function.
        return 0.5*erfc(-0.707106781186547524*(x-mu)/sig);
    }
    Doub invcdf(Doub p) {
    Return inverse cumulative distribution function.
        if (p <= 0. || p >= 1.) throw("bad p in Normaldist");
        return -1.41421356237309505*sig*inverfc(2.*p)+mu;
    }
};
```

Matlab and Mathematica both have many distributions, e.g.,

```
chi2pdf(x,v)
chi2cdf(x,v)
chi2inv(p,v)
```