Lecture 15, Feb 13, 2023

The Normal as a Limit of the Binomial

- Recall that for the binomial distribution $\mu = np, \sigma^2 = np(1-p)$
- Let X be distributed according to the binomial distribution, and let $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$
- As $n \to \infty$, Z approaches the standard normal

The Gamma Distribution

Definition

The Γ function is defined as:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, \mathrm{d}x, \alpha > 0$$

- Note $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$
  - Γ can be thought of as a continuous generalization of the factorial

Definition

The *gamma distribution* with parameters α, β is

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

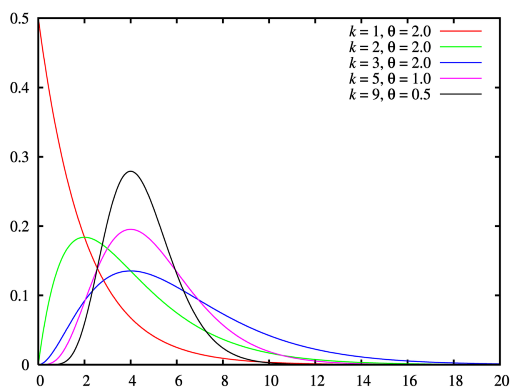

Figure 1: The gamma distribution ($k = \alpha$, $\theta = \beta$)

- Statistics:
  - $\mu = \alpha\beta$
  - $\sigma^2 = \alpha\beta^2$
- The gamma distribution combines and generalizes multiple distributions
- Note $\dfrac{1}{\beta^\alpha \Gamma(\alpha)}$ is just normalization (there is no x); the gamma function has no real effect on the shape of the distribution

The Chi-Squared Distribution

Definition

The χ^2 *distribution* is

$$f(x;v) = f\left(x;\alpha=\frac{v}{2},\beta=2\right) = \begin{cases} \frac{1}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)}x^{\frac{v}{2}-1}e^{-\frac{x}{2}} & x>0 \\ 0 & x\leq 0 \end{cases}$$

- This distribution is the distribution of the variance of random data

The Exponential Distribution

Definition

The *exponential distribution* is

$$f(x;\beta) = f(x;\alpha=1;\beta) = \begin{cases} \frac{1}{\beta}e^{-\frac{x}{\beta}} & x>0 \\ 0 & x\leq 0 \end{cases}$$

Where the random variable X is the time between events, given a mean time between events of $\beta = \frac{1}{r}$ where r is the mean rate of events

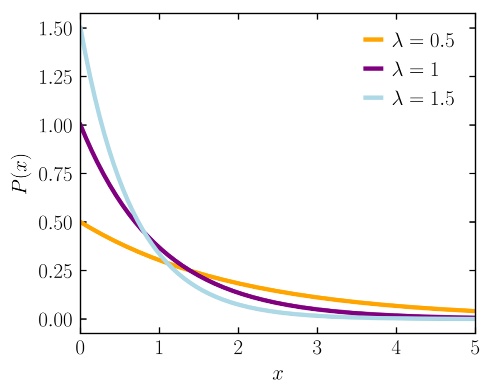

Figure 2: The exponential distribution ($\lambda = \beta$)

- Statistics:
  - $\mu = \beta$
  - $\sigma^2 = \beta^2$
- This is a decaying exponential which decays faster with larger β
  - Smaller values of β start off higher but decay more quickly
- This is the distribution of how long we need to wait for an event, given a mean waiting time of β
  - $\beta = \frac{1}{r}$ where r is the rate of events
  - Similar to a discrete version of the inverse binomial distribution
- Relation to the Poisson distribution: $p(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$
  - The Poisson distribution gives us the distribution for the number of events in an interval of length t where $\lambda = rt$

  - The exponential distribution gives us the time between events
  - The probability of no events occurring in t is $p(0; rt) = e^{-rt}$, which is also the probability that the first event occurs after time t
  - Let X be the random variable for time to the first event, then $P(X \geq x) = p(0; rx) = e^{-rx}$
  - From this we can get a CDF $F(x) = P(X \leq x) = 1 - P(X \geq x) = 1 - e^{-rx}$
  - The PDF is then $\frac{\mathrm{d}}{\mathrm{d}x}F(x) = re^{-rx} = f\left(x; \beta = \frac{1}{r}\right)$
- Example: on an average day a component fails every $\beta = 4$ days; if the failures are described by an exponential distribution, what is the chance a component lasts more than a week?
  - We want $P(X > 7) = 1 - P(X \leq 7)$
  - This is given by $\int_7^\infty \frac{1}{\beta}e^{-\frac{t}{\beta}}\,\mathrm{d}t = 0 - (-e^{-\frac{7}{4}}) = 0.17$
  - Note the exponential distribution is a continuous probability distribution so we used an integral
- Related question: how many failures should we expect in a week?
  - This is a Poisson distribution with $\lambda = rt = \frac{1}{\beta}t = \frac{1}{4} \cdot 7 = \frac{7}{4}$
  - Therefore the expected number of failures is just $\frac{7}{4}$