CHAPTER 6. COMMON CONTINUOUS DISTRIBUTIONS (AND HOW THEY ARISE)

Refer to table for pdf, cdf, mean, variance etc.

6.1 Uniform and uniform related

6.1.1 Uniform distribution

Distribution with constant PDF in a finite region .

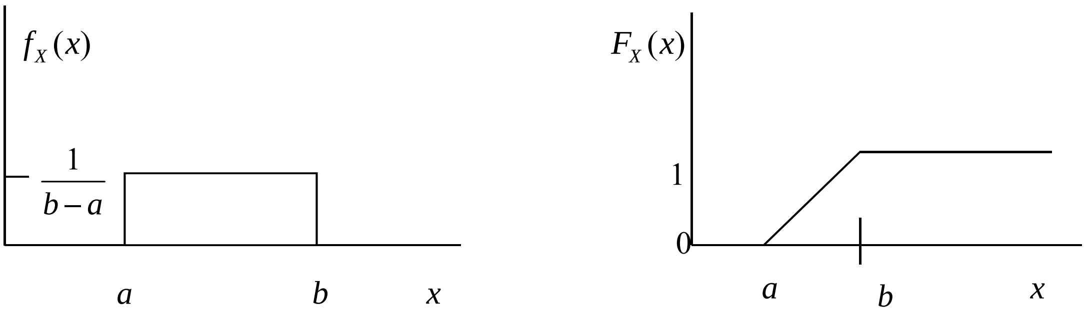


Figure 6-1: the PDF (left) and the CDF (right) of the uniform distribution

The PDF is:

$$f_X(x)=\begin{cases}\frac{1}{b-a}, & a \le x \le b\\ 0, & \text{elsewhere}\end{cases} \tag{6.1}$$

The CDF is:

$$F_X(x)=\begin{cases}0 & , \quad x<a\\ \frac{x-a}{b-a}, & a \le x \le b\\ 1, & x>b\end{cases} \tag{6.2}$$

The mean and variance are:

$$\mu = \frac{a+b}{2}$$
$$\sigma^2 = \frac{1}{12}(b-a)^2 \tag{6.3}$$

It is easy to check that odd central moments are zero. For example: $\mu_3 = 0$

The standard uniform distribution has $a = 0$ and $b = 1$.

<u>Applications</u>

1) Simulation of r.v.s.

2) For a r.v. when you know its range, but do not believe any interval is more likely than the other. e.g. , wind direction .

It can be shown that the Uniform distribution is the maximum entropy distribution given a set of lower and upper limits.

Its characteristic function is:

$$M_X(\theta) = \int_a^b e^{ix\theta} \frac{1}{b-a} dx = \frac{1}{b-a}\frac{1}{i\theta} e^{ix\theta/b} = \frac{e^{ib\theta} - e^{ia\theta}}{i\theta(b-a)} \tag{6.4}$$

6.1.2 Cauchy

Consider a non-linear but monotonic transformation of the uniform random variable Θ distributed between –π/2 and π/2:

$$\Theta \sim U(-\pi/2, \pi/2)$$
$$X = \tan\Theta \tag{6.5}$$

This transformation can be imagined to physically occur as follows: A laser source located at a fixed point A is used to illuminate a point on the wall (see Figure 6-2). The wall extends infinitely on the left and on the right. The location of point B depends on θ, the angle that the laser gun makes with the perpendicular. Since θ is random, so is *x*, and we denote them as Θ and *X* respectively. For simplicity, we take *l* = 1. We wish to determine the distribution of *X*.

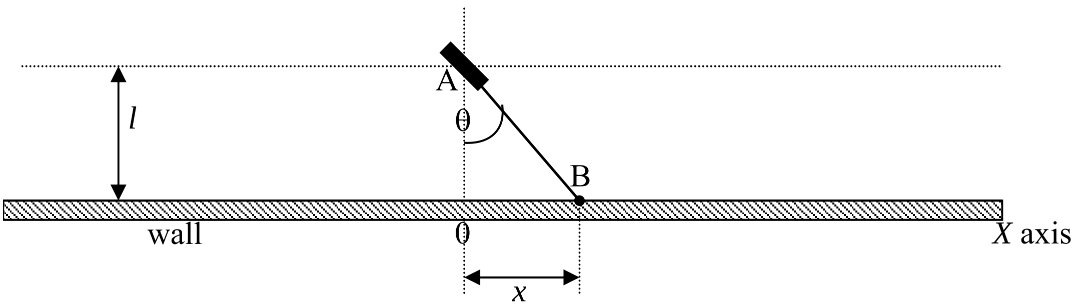


Figure 6-2: setup for the Cauchy distribution. The angle Θ is uniformly distributed, and its tangent *X* is the Cauchy random variable.

By definition,

$$F_X(x) = P[X \le x] = P[\tan\Theta \le x] = P[\Theta \le \tan^{-1} x] \tag{6.6}$$

We can write the last expression on the right since the transformation between Θ and *X* is monotonic. Now using the fact that $\Theta \sim U(-\pi/2, \pi/2)$, we obtain the CDF of *X* as:

$$F_X(x) = \frac{\tan^{-1} x - (-\pi/2)}{\pi/2 - (-\pi/2)} = \frac{\tan^{-1} x + \pi/2}{\pi}, \quad -\infty < x < \infty \tag{6.7}$$

which is the Cauchy distribution function. The density function of *X* is obtained by differentiation:

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < \infty \tag{6.8}$$

What about the mean and variance of *X*? It turns out that moments of any order do not exist for the Cauchy distribution. Recall the definition of the "improper integral" in calculus:

$$I = \int_{-\infty}^{\infty} g(x)dx = \lim_{l\to-\infty}\int_{l}^{c} g(x)dx + \lim_{l\to\infty}\int_{c}^{l} g(x)dx \tag{6.9}$$

for the integral *I* to exist, each of the two limits on the right must exist separately. When $g(x) = x/(1+x^2)$ it is easy to show that neither integral on the right is finite hence their sum does not exist. Figure 6-3 shows the distribution of the sample mean of samples chosen from the Cauchy distribution – for three different sample sizes 10^4, 10^5 and 10^6. Clearly, the law of large numbers does not work here and the estimate does not converge. Why?

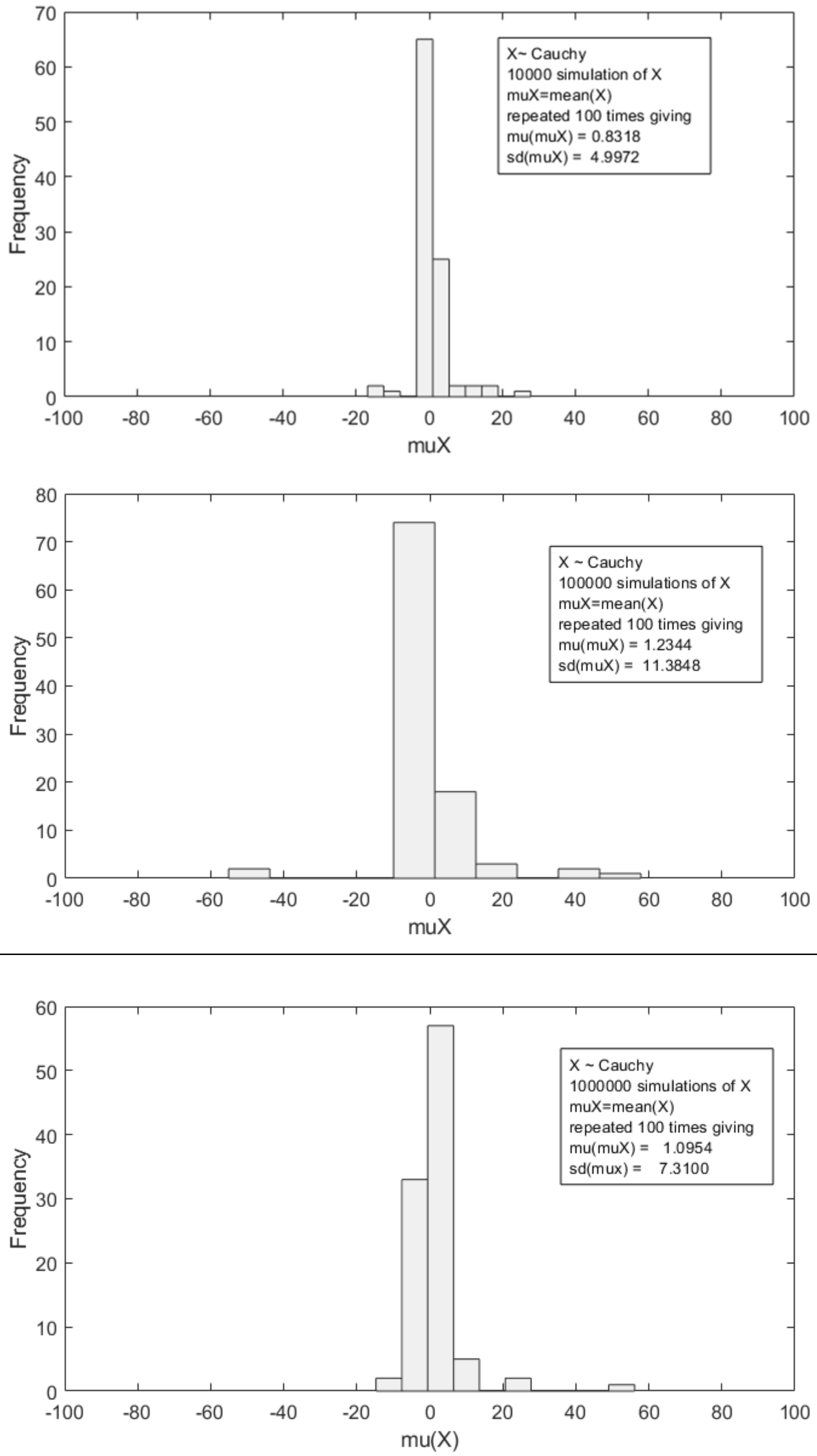


Figure 6-3: The mean does not exist for the Cauchy distriubtion. For three different sample sizes, the three panels show the distribution of the sample mean for IID samples drawn from a Cauchy distribution. The estimated mean does not converge.

6.2 Poisson process related: Exponential and Gamma (Erlang)

The Poisson process is discussed in detail in CHAPTER 11. It encompasses three distributions: the Poisson distribution which counts the random number of occurrences in a finite time interval, the exponential distribution which describes the time between successive occurences and the Erlang distribution which describes the time to a given integer number of occurrences. The Gamma distribution can be thought of as the generalization of the Erlang distribution. Of these, the Poisson random variable was discussed in CHAPTER 5. In the following subsections, we take up the continuous cases.

6.2.1 Exponential

The PDF and CDF of the exponential random variable are:

$$f_T(t) = \lambda e^{-\lambda t}, t \geq 0$$
$$F_T = 1 - e^{-\lambda t}, t \geq 0 \tag{6.10}$$

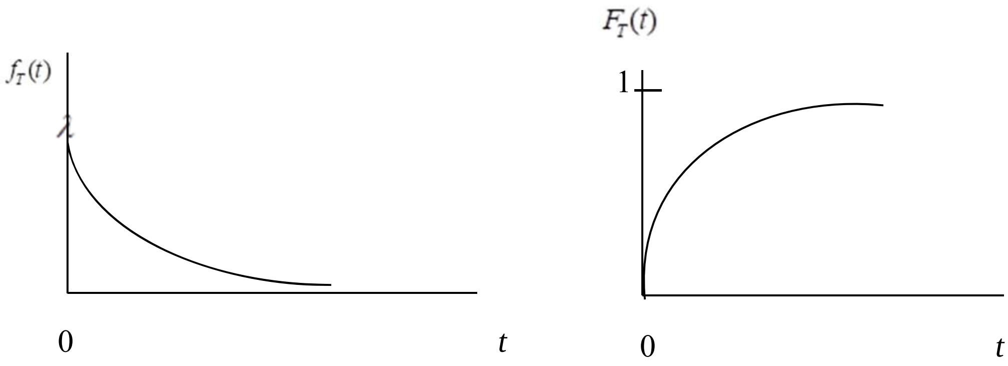


Figure 6-4: the PDF (left) and the CDF (right) of the exponential distribution

Its mean and variance are:

$$\mu_T = \frac{1}{\lambda}$$
$$\sigma_T^2 = 1/\lambda^2$$

The exponential distribution has the "memoryless" property:

$$P[T > t_0 + t \mid T > t_0] = P[T > t] \tag{6.11}$$

which is easy to prove.

The moment generating function is:

$$G(s)=E\left(e^{SX}\right)=\int_0^\infty e^{sx}\lambda e^{-\lambda x}dx$$
$$=\frac{\lambda}{\lambda-s};\quad s<\lambda$$

The shifted exponential distribution is useful for cases when there is a non-zero lower limit:

$$f_T(t)=\lambda e^{-\lambda(t-a)},\ t\ge a$$
$$F_T=1-e^{-\lambda(t-a)}\ ,\ t\ge a \qquad (6.12)$$

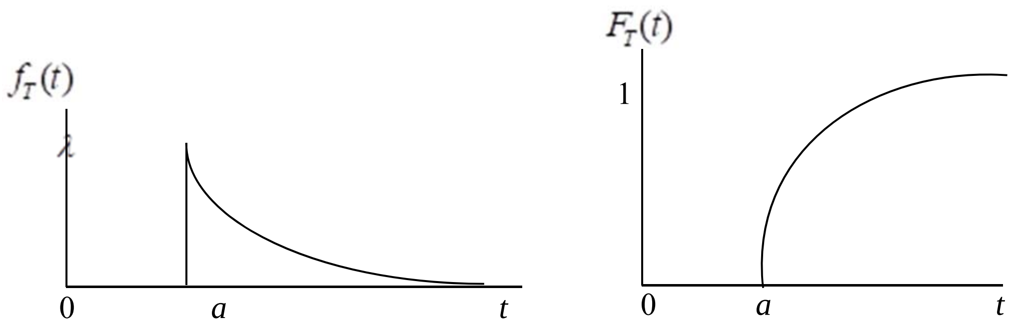


Figure 6-5: the PDF (left) and the CDF (right) of the shifted exponential distribution

6.2.1.1 Example: Exponential time to failure

The time to failure of a certain kind of industrial bulb is Exponential with mean 5 yrs. 10 bulbs are installed at a site. What is the probability that more than one bulbs are working after 8 yrs ? Bulb failures happen independently of one another.

10 trials, each with probability *p* of success.

$$p=P[T>8]=e^{-\lambda t}=e^{-8/5}=0.202$$

$$P[X>1]=1-P[X=0]-P(X=1)$$
$$=1-(1-p)^{10}-10p(1-P)^9$$
$$=1-.105-0.265$$
$$=0.63$$

6.2.1.2 Example: proof loading

The theoretical strength of a beam is exponentially distributed with mean 10 kNm. A moment of 12 kNm is applied on the beam (at the critical cross section) and it is found to survive without damage. What can you now say about the mean and standard deviation of the beam's strength?

6.2.2 Mean return period for continuous occurrence times

Let λ be the rate of a Poisson process. We are interested in the average time $\bar{T}$ between successive occurrences. As we will see in Section 11.2.1 that inter arrival times $\{T_i\}$ in a Poisson process are IID Exponentials with parameter λ . Then $\bar{T}$ is simply the average of T_i :

$$\bar{T} = 1/\lambda \tag{6.13}$$

6.2.2.1 10% in 50 year Earthquake

What is the mean return period of the earthquake that has a 10% probability of exceedance in 50 years?

Solution: Earthquakes occur according to a Poisson process with rate λ . We need to find λ such that:

$$P[N(50yr) \geq 1] = 0.10 \tag{6.14}$$

Hence,

$$\begin{aligned} &1 - P[N(50yr) = 0] = 0.10 \\ &or, 1 - \exp(-50yr\,\lambda) = 0.10 \\ &or, \lambda = 2.1072 \times 10^{-3} / yr \end{aligned} \tag{6.15}$$

Thus,

$$\bar{T} = 1/\lambda = 1/2.1072 \times 10^{-3}/yr = 474.6yr \tag{6.16}$$

Ans: approximately 475 years.

6.2.3 Erlang distribution

It has two parameters : $\lambda \& k$

Interpretation : RV describing time to kth arrival in a homogeneous Poisson process with rate λ .

$k = 1 \Rightarrow$ Exponential

$T_k = \tau_1 + \tau_2 + ... + \tau_k$, $\tau_i \sim \varepsilon(\lambda)$, and τ_i, τ_j indep for $i \neq j$

The CDF is given by the complementary CDF of the Poisson RV:

$$P(T_k \le t) = P(N(t) \ge k) = \sum_{n=k}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{\lambda !} \tag{6.17}$$

The PDF can be obtained by differentiating the CDF:

$$f_{T_k}(t) = \frac{1}{(k-1)!} \lambda e^{-\lambda t} (\lambda t)^{k-1},\ t \ge 0 \tag{6.18}$$

The first two moments are easily obtained since T_k is the sum of k iid RVs:

$$\mu = k/\lambda \qquad \sigma^2 = k/\lambda^2 \tag{6.19}$$

6.2.4 Gamma distribution

Generalization of Erlang when k is non integer. The density and distribution functions are:

$$f_X(x) = \frac{\lambda(\lambda x)^{k-1}}{\Gamma(k)} e^{-\lambda x},\ x \ge 0$$
$$F_X(x) = \frac{1}{\Gamma(k)} \int_0^x \lambda e^{-\lambda t} (\lambda t)^{k-1} dt = \frac{1}{\Gamma(k)} \int_0^{\lambda x} e^{-v} v^{k-1} dv,\ \lambda t = v \tag{6.20}$$

We call k the shape parameter, and $1/\lambda$ the scale parameter. The first two moments of the gamma random variable have the same form as for Erlang:

$$\mu = k/\lambda \qquad \sigma^2 = k/\lambda^2 \tag{6.21}$$

An important property of the gamma distribution is that the sum of independent gamma RVs is also gamma distributed (if rate remains the same):

$$\text{If } X_1 \sim Gamma(k_1, \lambda) \text{ and } X_2 \sim Gamma(k_2, \lambda) \text{ then, } Y = X_1 + X_2 \text{ is } Gamma(k_1 + k_2, \lambda) \tag{6.22}$$

6.2.5 Characteristic functions

6.2.5.1 Exponential

$$f_X(x) = \lambda e^{-\lambda x}, x > 0$$
$$M_X(\theta) = \left(1 - \frac{i\theta}{\lambda}\right)^{-1}$$

6.2.5.2 Gamma

gamma $f_X(x) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{k-1} e^{-\lambda x}, \quad x > 0, \ k, \lambda > 0$

$$M_X(\theta) = \left(1 - \frac{i\theta}{\lambda}\right)^{-k}$$

6.3 Normal and normal-related

The normal distribution is the most important and the most widely used distribution in probability and statistics. The N. D. arises naturally in many situations. It is also supported by central limit theorem – one of the most powerful theorems in probability theory. The normal density function is symmetrical about the mean:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty \quad (6.23)$$

where $\mu = \text{mean}, \sigma = \text{standard deviation}$. For the standard normal distribution, $\mu = 0, \sigma = 1$. The symbol Z is often used for the standard normal variable: Z~N(0,1). Its PDF is sketched in Figure 6-6. Note the symmetry about 0. The CDF of the standard normal variable, commonly denoted by Φ, is extensively tabulated.

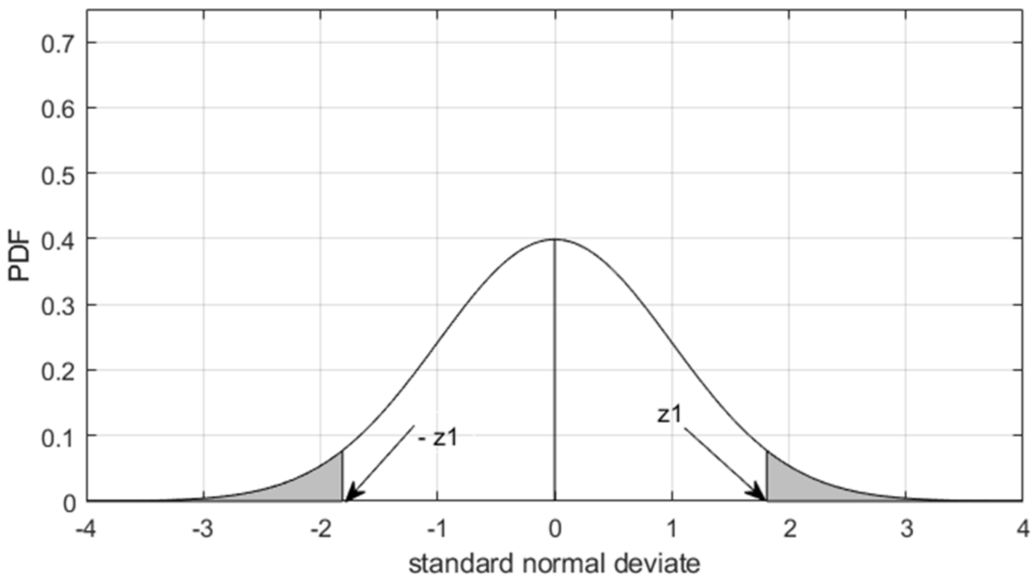


Figure 6-6: The standard normal PDF. Due to the symmetry about 0, the shaded area to the left of –z1 is equal to the shaded area to the right of +z1.

6.3.1 Central limit theorem.

The sum of *n* independent random variables $S_n = \sum_{k=1}^{n} X_k$, when centralized and normalized to zero mean & unit s.d., tends to the standard normal variable as $n \to \infty$ regardless of the distributions of the individual X_k's, provided three conditions are met :

1. $E[|X_k|] < \infty$ for all k .
2. $E[|X_k - E(X_k)|^{2+\delta}] < \infty$ for $\delta > 0$ and all k.
3. Lyapunov's condition:

$$\lim_{n\to\infty} \frac{1}{(\sigma_{S_n})^{2+\delta}} \sum_{k=1} E[|X_k - E(X_k)|^{2+\delta}] = 0$$

where $\delta > 0$ and $\sigma_{S_n}^2 = \sum_{k=1}^{n} \sigma_{X_k}^2$

For proof, refer to Lin p 68.

What is the implication of Lyapunov's condition?

Loosely speaking, CLT means:

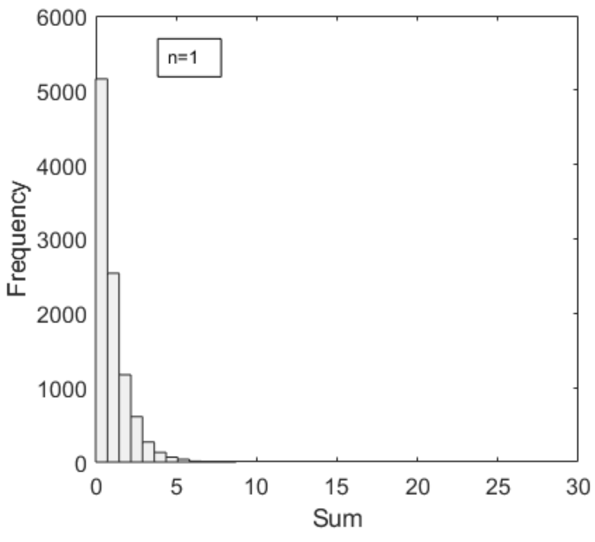
The sum of a large no of RVs,

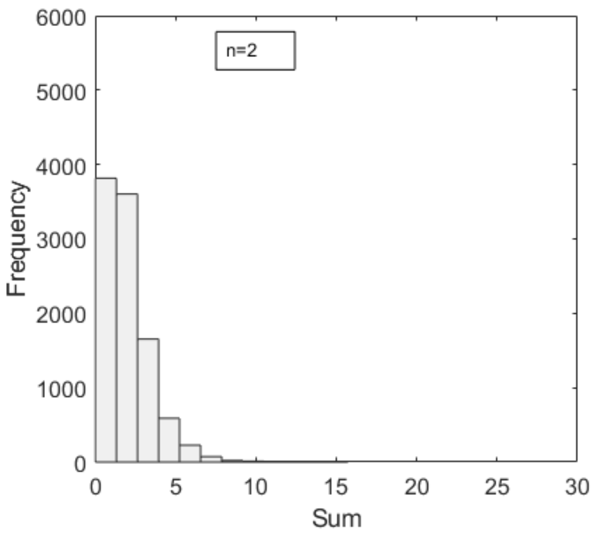
(i) without a single dominant components among them
(ii) without significant dependence among them

approaches the Normal RV regardless of the individual distributions.

If individuals are Normal, then sum is Normal regardless of (i) & (ii) above.

6.3.1.1 <u>Example: sum of IID exponentials</u>





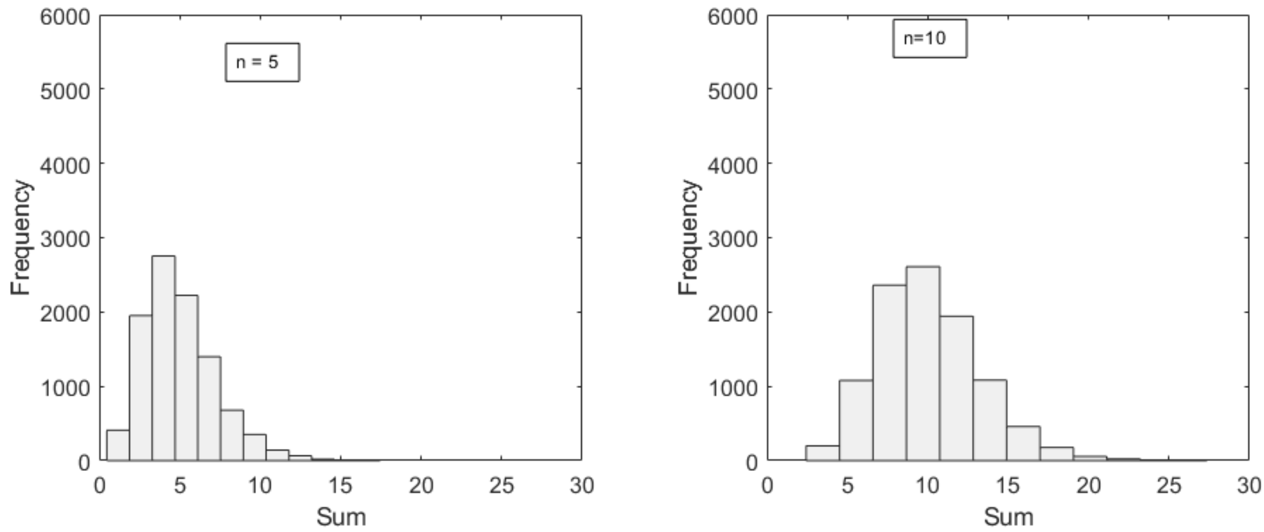


Figure 6-7: Illustration of CLT by summing *n* IID exponentials ($\lambda = 1$). Four values of *n* are considered with 10000 sequences each. Clearly, it does not take too many members for the sum to approach normality.

6.3.2 Linear transformation of Normal Variables

If $X \sim N(\mu_X, \sigma_X)$, then any linear transformation of *X*,

$$Y = aX + b \tag{6.24}$$

is also normal with mean and s.d. given by,

$$\mu_Y = a\mu_X + b, \qquad \sigma_Y = a\sigma_X \tag{6.25}$$

Therefore, the transformation of any normal variable $X \sim N(\mu_X, \sigma_X)$ that removes the mean and scales by the s.d.

$$Z = \frac{X - \mu_X}{\sigma_X} \tag{6.26}$$

yields the standard normal variable $(\mu = 0, \sigma = 1)$.

The CDF $F_Z(z)$ of the standard normal variable is commonly denoted by $\Phi(z)$ and the standard normal PDF by $\phi(z)$:

$$\begin{aligned} \phi_u(u) &= \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \\ \phi_u(u) &= \frac{d}{du}\Phi(u) \end{aligned} \tag{6.27}$$

6.3.3 Properties of the standard normal CDF

The importance of the standard normal CDF is that the CDF of any normal variable *X* can be obtained in terms of Φ as:

$$F_X(x) = P[Z < \frac{x-\mu_X}{\sigma_X}] = \Phi\left(\frac{x-\mu_X}{\sigma_X}\right), \; Z \sim N(0,1) \tag{6.28}$$

Specifically,

(i) $\Phi(-u) = 1 - \Phi(u)$

(ii) If $\Phi(u) = p$ then $u = \Phi^{-1}(p) = -\Phi^{-1}(1-p)$

(iii) $\Phi(-\infty) = 0, \Phi(0) = 1/2, \Phi(1) = .84, \Phi(2) = .977, \Phi(3) = .999$ etc.

6.3.4 Approximations to the standard normal CDF

The normal CDF is not available in closed form. However, over the years numerous closed form approximations to the standard normal CDF have been, and continue to be, proposed. Here we list a few well-known ones:

$$\begin{aligned}
\Phi_{Polya}(z) &= 0.5 + 0.5\sqrt{1-\exp(-2z^2/\pi)} \\
\Phi_{Tocher}(z) &= \frac{\exp\left(2z\sqrt{2/\pi}\right)}{1+\exp\left(2z\sqrt{2/\pi}\right)} \\
\Phi_{Norton}(z) &= \begin{cases} 1-0.5\exp\left(-\dfrac{z^2+1.2z^{0.8}}{2}\right), 0 \le z \le 2.7 \\ 1-\dfrac{\phi(z)}{z}, \quad z > 2.7 \end{cases} \\
\Phi_{Lin}(z) &= 1-0.5\exp\left(-0.717z-0.416z^2\right)
\end{aligned} \tag{6.29}$$

Figure 6-8 shows how these approximations perform related to the exact function.

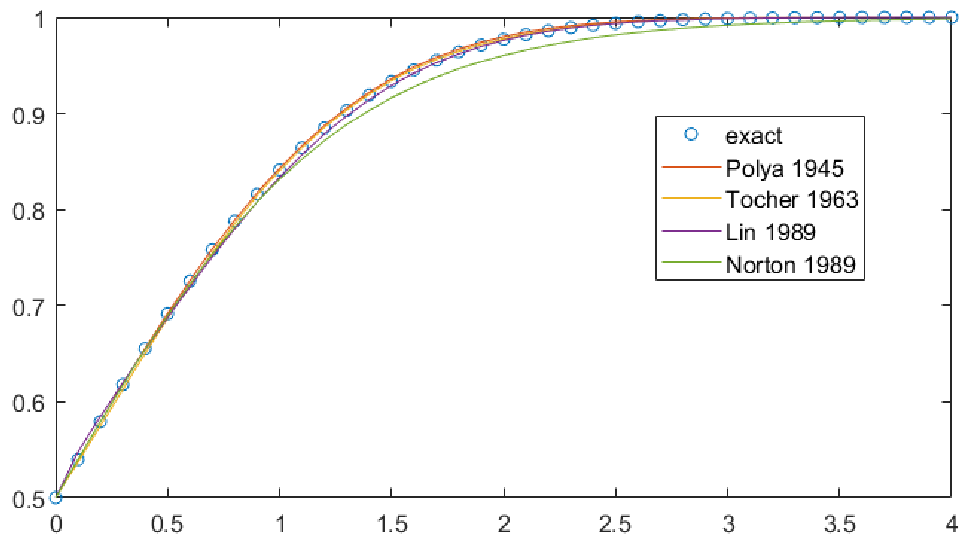


Figure 6-8: Approximations to the standard normal distribution function. Due to symmetry, only positive z values are shown.

6.3.5 *Examples:*

1. A batch of bearings of nominally identical diameter are tested for acceptance. It is known that the bearing diameter, D, is Normally distributed with mean 5 mm and s. d. 0.1 mm. The acceptance criteria is that D has to be within the range 5 mm   d. What is the value of the tolerance, d, such that only 10% of the bearings are rejected?

Ans: $d = 1.65 \times .1\,\text{mm} = 0.165\,\text{mm}$.

2. A soft drink machine is regulated so that it dispenses an average of 7 ounces per cup. If the amount dispensed is Normally distributed with standard deviation 0.5 ounce,

a) how many 8-ounce cups will likely overflow out of the next 1000 drinks ?

b) what fraction of cups will have between 6.7 and 7.3 ounces?

Solution:

$a)$ $X =$ amount dispensed.

$$P[X > 8] = P[U > \frac{8-7}{.5}] = P[U > 2] = 1 - \Phi(2) = 1 - .977 = .023$$

This is a Binomial problem with $n = 1000$, $p = 0.023$. Hence, the expected no. of cups to overflow = 23, and the SD is 4.7.

b) $$P[6.7 < X \le 7.3] = \Phi\left(\frac{7.3-7}{.5}\right) - \Phi\left(\frac{6.7-7}{.5}\right) = \Phi(.6) - \Phi(-.6) = 2\Phi(.6) - 1 = 0.45.$$

3. In a statistics course, the lowest ten percent scores were given the F grade. The mean score was 74 and the standard deviation was 7.9. Find the cutoff score for passing the course.

Ans: 63.9

4. The average life of a type of fan motor is 10 years with a standard deviation of 2 years. The manufacturer wants to come up with a warranty policy so that only 3% of the motors need to be replaced.

(a) What should the warranty period be if the motor life follows the normal distribution.

(b) What if the motor life follows the exponential distribution?

(c) What if the motor life follows the uniform distribution?

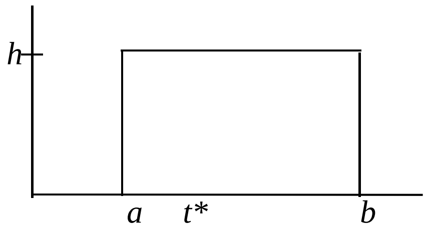
Ans:

(a) $P[T < t^*] = .03$, $t^* = 10 - 2\Phi^{-1}(.97) = 10 - 2\times 1.88 = 6.24$ yr.

(b) Since the Exponential is a one-parameter distribution, ignore the higher order moment, and use the mean to obtain the parameter: $\mu = 10yr \Rightarrow \lambda = 1/10yr$

$$P[T < t^*] = 1 - e^{-\lambda t^*} = 0.03 \Rightarrow t^* = 3\,\text{yr}$$

(c)



$$(a+b)/2 = 10, (b-a)^2/12 = 2^2 = 4$$
$$\Rightarrow a = 6.5, b = 13.5, h = 1/7$$

Required t^* such that $(t^* - a)h = .03$

$$\Rightarrow t^* = 6.7\,\text{yr}$$

6.3.6 Example: six sigma methodology

In a production process, suppose the acceptance criteria is two-sided. This means there is a certain measurable property, X, of the product. The target value for X is τ, and there is a tolerance of $\pm\Delta$ about τ such that the upper specification limit is $\tau + \Delta$, and the lower specification limit is $\tau - \Delta$. Due to various types of variations in the production process, the property X is generally described as a random variable – its mean is μ, standard deviation is σ and X is commonly assumed to be normally distributed. An item will be "out of specification" and be rejected if it exceeds the tolerance:

$$F = \{\text{item is out of specification}\} = \{X < \tau\text{-}\Delta \cup \tau + \Delta < X\} \qquad (6.30)$$

Quality control can ensure that the fraction of rejected products remains acceptably low.

In the short term, the mean of X is kept on target. The probability of rejection in the short term is therefore:

$$p_s = P\{X < \tau\text{-}\Delta \cup \tau+\Delta < X\} = \Phi\left(\frac{-\Delta}{\sigma}\right) + 1 - \Phi\left(\frac{\Delta}{\sigma}\right) = 2\Phi\left(-\frac{\Delta}{\sigma}\right) \tag{6.31}$$

The variability in production is captured by the standard deviation σ, and a small enough σ can ensure that p_s is acceptably small. "Three sigma control" strives to limit σ to Δ/3, so that the short term probability of being out of specification is $2\Phi(-3) = 0.0027$. In other words, there will be 2700 defects per million (DPM) under three sigma control.

Although the mean of X can initially be set on the target, the distribution of X tends to drift away from the target in the long term. Say, a drift by as much as 1.5σ is allowed. Under three sigma control, the long term probability of rejection becomes:

$$\begin{aligned} p_l &= P\{X < \tau\text{-}\Delta \cup \tau+\Delta < X\} = \Phi\left(\frac{\tau-\Delta-(\tau+1.5\sigma)}{\sigma}\right) + 1 - \Phi\left(\frac{\tau+\Delta-(\tau+1.5\sigma)}{\sigma}\right) \\ &= \Phi\left(-\frac{3\sigma+1.5\sigma}{\sigma}\right) + \Phi\left(-\frac{3\sigma-1.5\sigma}{\sigma}\right) \\ &= \Phi(-4.5) + \Phi(-1.5) \\ &= 0.0668 \end{aligned} \tag{6.32}$$

That is, under three sigma control in the presence of long term drift, the DPM value is 66800 which is very high even for a single item.

With increasing complexity of manufactured systems, this long-term probability of being out of spec of a single item can lead to unacceptably high rejection of products with a large number of elements. Six sigma control strives to ensure tighter quality control by limiting σ to Δ/6. This brings down p_l to:

$$\begin{aligned} p_l &= \Phi\left(-\frac{6\sigma+1.5\sigma}{\sigma}\right) + \Phi\left(-\frac{6\sigma-1.5\sigma}{\sigma}\right) \\ &= \Phi(-7.5) + \Phi(-4.5) \\ &= 3.340\,\mathrm{e}-6 \end{aligned} \tag{6.33}$$

that is, the DPM number comes down from 66800 to 3.34.

6.3.7 *The challenge of course is being able to limit σ to Δ/6.*

6.3.8 *Characterstic function of the normal random variable*

$$M_{X(\theta)} = \int_{-\infty}^{\infty} e^{i\theta x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Consider the exponent of the integrand:

$$-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+i\theta x=-\frac{x^2}{2\sigma^2}+\frac{x\mu}{\sigma^2}-\frac{\mu^2}{2\sigma^2}+i\theta x$$
$$=-\frac{1}{2\sigma^2}\left(x^2-2x\mu+\mu^2-2\mathrm{i}\theta\sigma^2 x\right)$$
$$=-\frac{1}{2\sigma^2}\left[x^2-2\left(\mu+i\theta\sigma^2\right)x+\left(\mu+i\theta\sigma^2\right)^2-\mu^2-2i\theta\mu\sigma^2-i^2\theta^2\sigma^4+\mu^2\right]$$
$$=-\frac{1}{2\sigma^2}\left[x^2-2\left(\mu+i\theta\sigma^2\right)x+\left(\mu+i\theta\sigma^2\right)^2+\theta^2\sigma^4-2i\theta\mu\sigma^2\right]$$
$$=-\frac{1}{2\sigma^2}\left[\left(x-\mu-i\theta\sigma^2\right)^2+\theta^2\sigma^4-2i\theta\mu\sigma^2\right]$$
$$\therefore M_X(\theta)=e^{-\frac{1}{2\sigma^2}\left[\theta^2\sigma^4-2i\theta\mu\sigma^2\right]}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left[\frac{x-\mu-i\theta\sigma^2}{\sigma}\right]^2}dx$$
$$=e^{-\frac{1}{2}\left[\theta^2\sigma^2-2i\mu\sigma\right]}\times 1$$
$$=e^{\left[i\theta\mu-\sigma^2\theta^2/2\right]}$$

6.3.9 M.G.F. for normal distribution

The moment generating function for the normal random variable with mean μ and variance σ^2 is:

$$G(S)=E\left(e^{SX}\right)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp(sx)\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$$
$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2-2\mu x-2s\sigma^2 x+\mu^2}{2\sigma^2}\right)dx$$

which can be simplified by rewriting the exponent as:

$$G(S)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2-2(\mu+s\sigma^2)x+(\mu+s\sigma^2)+(2\mu s\sigma^2+s^2\sigma^4)}{2\sigma^2}\right)dx$$
$$=\exp\left(\mu s+s^2\sigma^2/2\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu-s\sigma^2)^2}{2\sigma^2}\right)dx$$

The integrand is the normal pdf and thus integrates to one, yielding:

$$G(s)=\exp\left(\mu s+s^2\sigma^2/2\right) \tag{6.34}$$

6.3.10 Higher moments

The normal distribution is completely defined by its first two moments. Higher moments can be given by differenting the MGF (Eq etc) as:

$$G^{'}(s)=\left(\mu+s\sigma^2\right)G(s)\Rightarrow E(X)=G^{'}(0)=\mu\times 1 \tag{6.35}$$

$$G''(s) = \sigma^2 G(s) + \left(\mu + s\sigma^2\right)^2 G(s) \Rightarrow E(X^2) = G''(0) = \sigma^2 + \mu^2 \tag{6.36}$$

$$\begin{aligned} G'''(s) &= \sigma^2\left(\mu + s\sigma^2\right)G(s) + 2\left(\mu + s\sigma^2\right)\sigma^2 G(s) + \left(\mu + s\sigma^2\right)^3 G(s) \\ &= \left[3\mu\sigma^2 + 3s\sigma^4 + \left(\mu + s\sigma^2\right)^3\right]G(s) \\ &\Rightarrow E(X^3) = G'''(0) = \mu^3 + 3\mu\sigma^2 \end{aligned} \tag{6.37}$$

$$\begin{aligned} G''''(s) &= \left[3\sigma^4 + 3\left(\mu + s\sigma^2\right)^2\sigma^2\right]G(s) + \left[3\mu\sigma^2 + 3s\sigma^4 + \left(\mu + s\sigma^2\right)^3\right]\left(\mu + s\sigma^2\right)G(s) \\ &\Rightarrow E(X^4) = G''''(0) = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4 \end{aligned} \tag{6.38}$$

and so on. We see that for the standard normal variable (i.e., $\mu = 0$ and $\sigma^2 = 1$):

$$E(X) = 0, E(X^2) = 1, E(X^3) = 0, E(X^4) = 3, \text{ and so on} \tag{6.39}$$

6.3.11 Sum of Normals

Let Y be a linear combination of a vector of normal variables $\underline{X} \sim N(\underline{\mu}_X, \underline{V}_X)$:

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i \tag{6.40}$$

where $\underline{\mu}_X, \underline{V}_X$ are respectively the mean vector and covariance matrix of $\underline{X}$. Then Y is normally distributed and the mean and variance of Y are:

$$\begin{aligned} \mu_Y &= a_0 + \underline{a}^T \underline{\mu_X} \\ \sigma_Y &= \underline{a}^T \underline{V}_X \underline{a} \end{aligned} \tag{6.41}$$

6.3.12 Exponentiated normal – the Lognormal distribution

Like the normal random variable is for the sum, the lognormal is the limiting case of the product of a large number of independent RVs.

Y is a lognormal RV means $X = \ln Y$ is normally distributed. Conversely, if X is Normal, then its exponential, $Y = e^X$ is lognormal. The first two moments of X and Y are related as follows:

$$\begin{aligned} &Y = e^X,\ X \sim N(\mu_X, \sigma_X) \text{ and } Y \sim LN(\mu_Y, \sigma_Y) \\ &\mu_X = \mu_{\ln Y} = \ln m_Y = \ln \mu_Y - \frac{1}{2}\sigma_{\ln Y}^2 \\ &\sigma_X = \sigma_{\ln Y} = \sqrt{\ln(1 + V_Y^2)} \end{aligned} \tag{6.42}$$

where $V_Y = \dfrac{\sigma_Y}{\mu_Y}$, and $m_Y =$ median of Y. Conversely,

$$\mu_Y = \exp\left(\mu_X + \sigma_X^2 / 2\right)$$
$$\sigma_Y{}^2 = (e^{\sigma_X^2} - 1)\mu_Y{}^2 \quad (6.43)$$
$$V_Y{}^2 = \exp(\sigma_X^2) - 1$$

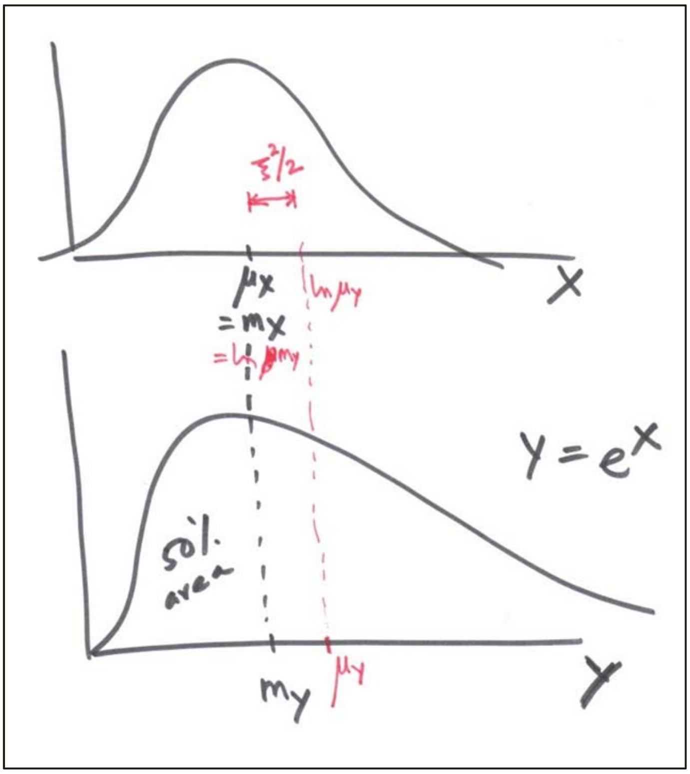


Figure 6-9: Normal to log normal transformation. Note that the median follows same transformation: $m_Y = \exp(m_X) = \exp(\mu_X)$

The CDF of *Y* is evaluated with the help of the corresponding normal parameters:

$$F_Y(y) = P[Y \le y] = P[\ln Y \le \ln y] = \Phi\left[\frac{\ln y - \mu_X}{\sigma_X}\right], y \ge 0 \quad (6.44)$$

The PDF of the lognormal random variable can be derived from the normal CDF as show in Section .

6.3.12.1 Example:

Y = Fatigue life of A 285 steel , Y ~ LN . The mean and sd of Y are:

$\mu_Y = 430000$ cycles and $\sigma_Y = 215000$ cycles under some loading .

Find the probability that life exceeds one million cycles.

$\Rightarrow V_Y = .5$, $\lambda = 12.86, \xi = .472$

$P[Y > 10^6] = 1 - \Phi(2.03) = .021$

6.3.13 Product of Lognormals

The lognormal family is closed under multiplication.

$Y = \alpha_0 Y_1^{\alpha_1} Y_2^{\alpha_2} ... Y_n^{\alpha_n}$

If $Y_i \sim LN$

$\Rightarrow Y \sim LN$

6.3.14 Sum of squared iid Normals – Chi-squared

If Z_1, Z_2, ..., Z_n are independent standard normal rvs, then $V = \sum_{i=1}^{n} Z_i^2$ is Chi-squared distributed with *n* dofs. The mean and variance are:

$$\mu = k\ ,\ \sigma^2 = 2k \tag{6.45}$$

Chi-squared distribution also arises as the limiting distribution of the sum, $X = \sum_{i=1}^{k} (N_i - np_i)^2 / (np_i)$, (how many dofs??? since there are two constraints: sum of N_i = *n* and sum of p_i=1) where N_1, N_2,..., N_k have a joint multinomial distribution with parameters *n*, p_1, p_2, ..., p_k. The gamma distribution with shape parameter *n*/2 and scale parameter 2 is the same as the chi-squared dist with *n* dofs.

6.3.15 Ratio of Normal and Chi-squared – the t-distribution

If standard normal *Z* and chi-squared *V* with dof *k* are independent of each other, then

$$T = \frac{Z}{\sqrt{V/k}}, \qquad Z \sim N(0,1),\ \ V \sim \chi^2(k) \tag{6.46}$$

has the t distribution with *k* d.o.f. Its first two moments are:

$$\begin{aligned} \mu &= 0 \\ \sigma^2 &= \frac{k}{k-2}(k > 2) \end{aligned} \tag{6.47}$$

The *t* distribution finds use in interval estimation of the mean when the population variance is unknown.

6.3.16 Ratio of two independent Chi-squared variables - F distribution

If W and Y are two independent Chi-squared random variables with dof's u and v, then,

$$F = \frac{W/u}{Y/v},\ W \sim \chi^2(u), Y \sim \chi^2(v) \tag{6.48}$$

has the F distribution. Its first two moments are:

$$\mu = \frac{v}{V-2} \qquad V = 2$$

$$\sigma^2 = \frac{2V^2(u+v-2)}{u(v-2)^2(v-4)} \qquad v > 4.$$

Where does it find use?

6.3.17 Wald (Inverse Gaussian)

In a Brownian motion with constant drift, *v*, and diffusion β, the distance travelled in a given time *t* is *X*. If the initial condition is non-random, it can be shown that *X* is Gaussian. If we flip the question, and ask for the time *T* required to reach (for the first time) a fixed distance *d*, then what is the distribution of *T*? The distribution of *T* can be shown to be Wald.

6.3.18 Maxwell Boltzmann distribution

Distribution of the speed of a particle when each component of velocity is zero-mean normally distributed and the components are mutually independent.

6.3.19 Normal as the limiting form

6.3.19.1 Normal approximation to Poisson

The Poisson CDF also approaches the normal CDF as long as $\mu \gg 1$.

Poisson $(\mu) \rightarrow N(\mu, \sqrt{\mu})$

The proof invokes CLT. Since the Erlang random variable is the sum of *k* IID Geometric random variables, it approaches the normal as *k* grows large. Since the Erlang CDF is 1 minus the Poisson CDF. Alternately, one can also describe the Poisson count in a given interval as the sum of counts from a partition of that interval since occurences in disjoint intervals are mutually independent.

Application of Stirling's formula also gives the same result.

6.3.19.2 Normal approximation to Binomial

Let $X \sim B(n, p)$ be binomially distributed. Define the normalized random variable, $Y = \dfrac{X - np}{\sqrt{npq}}$. It can be shown that: $Y \rightarrow Z$ as $n \rightarrow \infty$ and $\min(np, nq) >> 1$ where $Z \sim N(0,1)$.

This follows directly from the CLT as *X* is the sum of *n* independent IID Bernoulli RVs.

6.3.19.3 Normal approximation to Gamma

etc.

6.3.19.4 Normal approximation to Chisq

etc.

6.4 Logistic

etc.

6.5 Pareto

The Pareto random variable has a density function that, like the Exponential, has its maximum at the left end point and drops monotonically with increasing *x*. The important diference is that the drop is less steep than the exponential and is called "heavy tailed" in comparison. The PDF is of the form:

$$f_X(x) = \alpha x^{-\beta},\ x > x_0 \tag{6.49}$$

with $\alpha > 0, \beta > 1$. Since the area under the PDF must equal 1, x_0 is related to α and β through $x_0^{\beta-1} = \alpha / (\beta - 1)$. Its CDF is given by:

$$F_X(x) = 1 - \frac{\alpha}{\beta - 1} x^{-(\beta-1)},\ x > x_0 \tag{6.50}$$

6.6 Models of extremes

The family of extreme value distributions, discussed in CHAPTER 9, arise as the asymtotic distribution for extremes of sequences observed in a wide class of natural and engineering phenomena. There are three asymptotic forms: the Gumbel (type 1), Frechet (type 2) and Weibull (type 3); each type lends itself to both minima and maxima of the sequence. Nevertheless, it is more common to talk about the Gumbel distribution for maxima, the Frechet distribution for maxima and the Weibull distribution for minima. The type of the extreme value distribution (for both maxima and minima) depends on the shape of the "parent" distribution, i.e., the distribution of the individual members of the IID sequence. For example if the parent distribution is normal, then the maximum of that sequence must be asymptotically type 1.

6.6.1 The Gumbel distribution for maxima

The Gumbel is a two parameter distribution. Its CDF is:

$$F_X(x) = e^{-e^{-\alpha(x-u)}},\ -\infty < x < \infty \tag{6.51}$$

The PDF is obtain by differentiating the CDF:

$$f_X(x) = \frac{d}{dx} F_X(x) = \alpha e^{-\alpha(x-u)} e^{-e^{-\alpha(x-u)}},\ -\infty < x < \infty \tag{6.52}$$

The parameter u is actually the mode of the distribution which can be verified from the stationarity condition $\frac{d}{dx} f_X(x) = 0$ and the sign of the second derivative.

The mean and SD are:

$\mu_X = u + \frac{\upsilon}{\alpha}$ where, $\upsilon = .5772$ (Euler's const.)

$$\sigma_X = \frac{\pi}{\sqrt{6}\alpha}$$

The median can be obtained by setting $F_X(x_m) = 0.5$:

$$x_m = u + 0.3665/\alpha$$

6.6.1.1 <u>Example</u>

Annual max wind speed is Gumbel distributed with mean 50 mph and cov 25 %. Find the 100 yr wind.

Solution:

Given, $\mu = 50, V = .25 \Rightarrow \sigma = 12.5$ we derive u and α:

$$\alpha = \frac{\pi}{\sqrt{6} \times 12.5} = 0.1026$$

$$u = \mu - \frac{\gamma}{\alpha} = 50 - \frac{.5772}{\alpha} = 44.4 \text{ mph}$$

Let x_{100} be the 100 year wind, i.e., $P[X > x_{100}] = 1/100$. Hence,

$$F_X(x_{100}) = 0.99$$

$$\Rightarrow x_{100} = u - \frac{1}{\alpha}\ln(-\ln .99) = u + \frac{4.600}{\alpha} = 89.2 \text{ mph}$$

6.6.2 Frechet Distribution for maxima

The Frechet distribution for maxima is limited on the left:

$$F_X(x) = e^{-\alpha(x-\lambda)^{-\gamma}}, \quad \gamma, \alpha > 0, x > \lambda \tag{6.53}$$

The logarithm of the two-parameter (λ=0) Frechet for maxima gives the Gumbel random variable. In case the Frechet distribution is used to model minima, it should be remembered that the Frechet distribution for minima is limited on the right.

6.6.3 Weibull Distribution for minima

The Type 3 extreme value distribution for smallest values with two parameters is one of the most important in engineering. The shape of its PDF depends on k:

$$f_Z(z) = \frac{k}{k}\left(\frac{z}{u}\right)^{k-1} \exp\left[-\left(\frac{z}{u}\right)^{k}\right],\ z \geq 0 \tag{6.54}$$

The CDF is:

$$F_Z(z) = 1 - \exp\left[-\left(\frac{z}{u}\right)^{k}\right],\ z \geq 0 \tag{6.55}$$

The mean, variance and COV are given by:

$$\begin{aligned} \mu_Z &= u\Gamma\left(1+\frac{1}{k}\right) \\ \sigma_Z{}^2 &= u^2\left[\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)\right] \\ \Rightarrow V_Z &= \sqrt{\frac{\Gamma\left(1+2/k\right)}{\Gamma^2\left(1+1/k\right)} - 1} \end{aligned} \tag{6.56}$$

Application : Modeling of the time to failure distribution. In case the Weibull distribution is used to model maxima, it should be remembered that the Weibull distribution for maxima is limited on the right.

6.6.3.1 Weibull approximation to Rayleigh, Exponential etc.

k = 1 $\Rightarrow$ Exponential Distribution.

The exponential random variable is a special case of the two parameter Weibull: k = 1. Its mean is equal to its mode, and its COV is 100%.

k = 2 $\Rightarrow$ Rayleigh Distribution

The Rayleigh random variable is a special case of the two parameter Weibull: k = 2. Its mean is 88.62% of its scale parameter, and its COV is 52.27%.

6.6.3.2 Examples:

Example 1:

Consider a structural steel cable for a bridge whose strength is Weibull distributed with mean = 300 kips and cov = 15%. The load is 208 kips. What is the probability of failure of the cable?

Given: V= 15% $\Rightarrow$ k = 8

u = $u = \mu / \Gamma(1+1/8) = 300/0.94 = 318.6$ kips.

Required, P[failure]=P[capacity < load] = ?

Solution:

$$P\left[X \le 208\,\text{kips}\right] = 1 - \exp\left[(-\left(\frac{208}{318}\right)^{8}\right] = 0.03$$

Example 2:

Suppose the life of emergency brake-sets in elevators is modeled as a Weibull random variable with mean 15 years and standard deviation 3 years. In order to make the system redundant, there are two sets of brakes in an elevator, and the elevator is safe as long as at least one set is OK. Failure of the brakes is independent of one another.

A new elevator is installed (with two sets of new brakes). You are asked to schedule the next maintenance time, t_0 (in years), such that braking system failure probability does not exceed 0.001 at the time of maintenance. Find t_0.