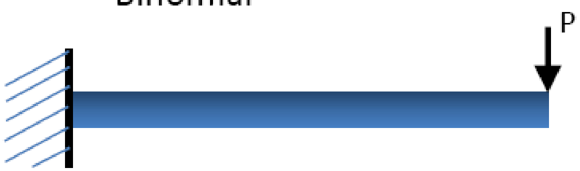
# Selection of Probability Distributions

- Several types of probability distributions are available to ***characterize uncertain system parameters***

- Selecting ***suitable probability distributions*** for system parameters is essential and largely depends on
  - Nature of structural system
  - Underlying assumptions of probability distributions
  - Distribution shape
  - Convenience and simplicity for subsequent computation

- ***Types of probability distributions*** most commonly used in engineering
  - Gaussian
  - Lognormal
  - Gamma
  - Binomial
  - Uniform
  - Extreme value
  - Exponential



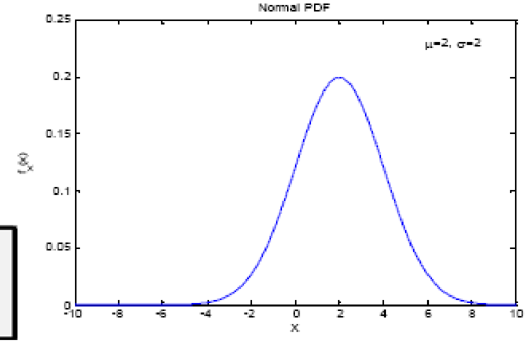

Tip displacement may largely depend on what type of distribution is assumed to characterize uncertainty in load P

# Common Engineering Distributions



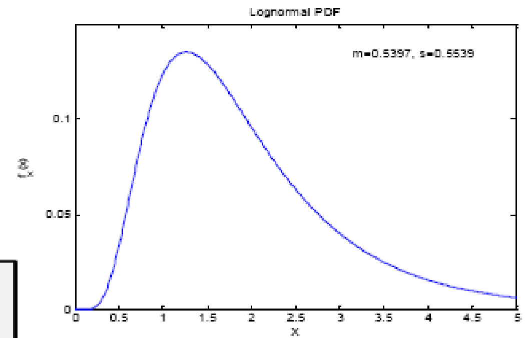
- Gaussian Distribution (Most common overall)
  - a.k.a. Normal distribution or “bell curve”
  - 2 parameter distribution
  - symmetric distribution with x from -∞ to ∞

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



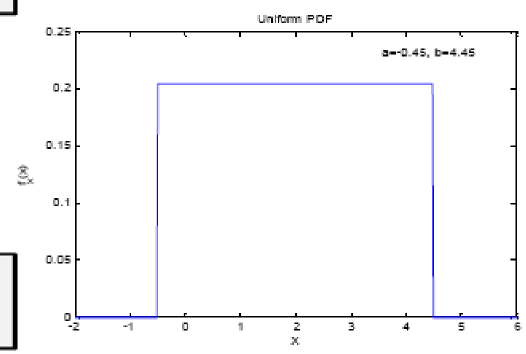

- Lognormal Distribution (Related to Normal)
  - If the *ln(x)* produces a normal random variable, then *X* is lognormally distributed
  - 2 parameter distribution
  - Parameters are not equal to mean and standard deviation

$$f_x(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left(-\frac{(lnx-m)^2}{2s^2}\right)$$




- Uniform Distribution (Simplest)
  - Simplest of continuous distributions
  - 2 parameter distribution
  - All positive values between *a* and *b* are equally likely

$$f_x(x) = \frac{1}{b-a}$$

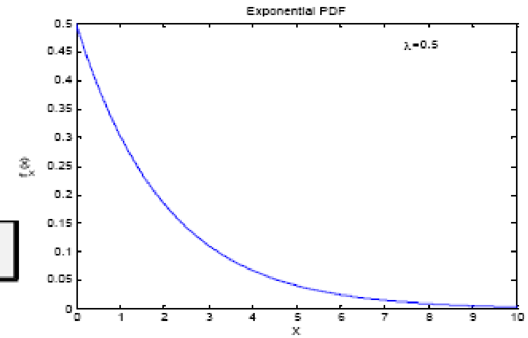


# Common Engineering Distributions



- Exponential Distribution (Single Parameter)
  - Simplest of continuous distributions that has only one parameter
  - Common in reliability Analysis: used for constant failure rates

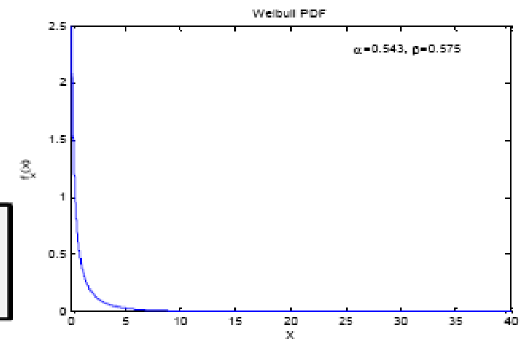
$$f_x(x) = \lambda exp(-\lambda x)$$




- Weibull Distribution (Fatigue of Materials)
  - Widely used in reliebiliy engineering: fatigue
  - Weakest link applications
  - Member of extreme value family

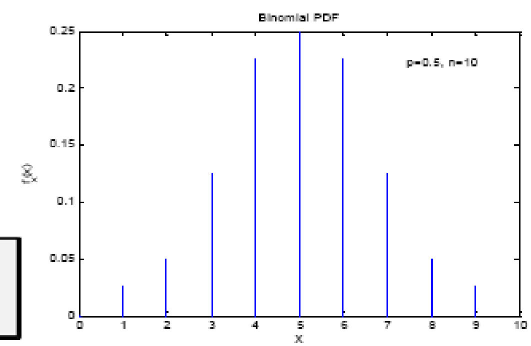
$$\alpha, \beta > 0$$

$$f_x(x) = \left(\frac{\alpha}{\beta}\right)\frac{x^{\alpha-1}}{\beta} exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right]$$




- Binomial Distribution (Counting process)
  - Generalization of Bernoulli for n independent trials
  - Gives probability of success of r unfavorable outcomes in n trials

$$f_x(x) = P(X = r) = \binom{n}{r} p^r (1-p)^{n-r} \qquad \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

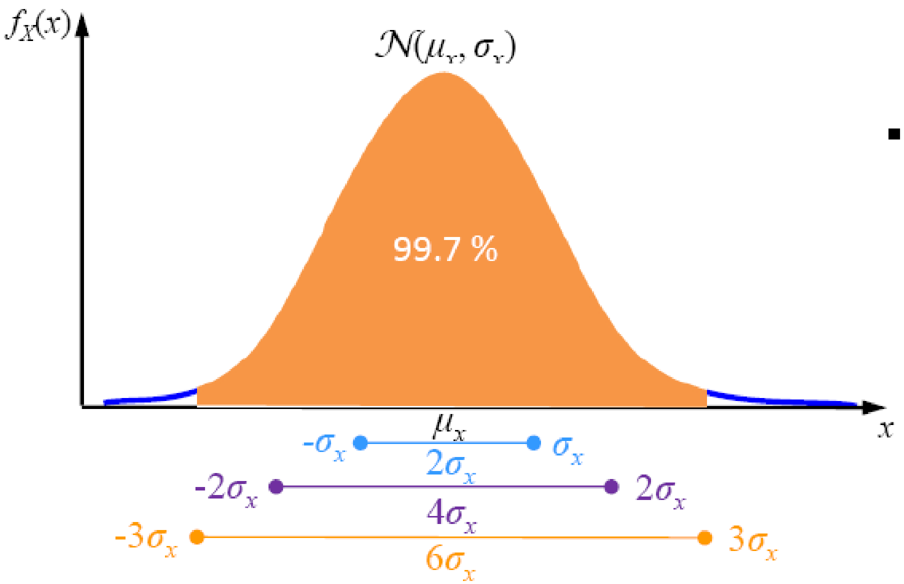


# Normal Distribution

## Normal (or Gaussian ) Distribution

- Used in many engineering and science applications due to its ***simplicity and convenience***

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right],\ -\infty < x < \infty$$




- Often used for problems involving small variation
  - Young's modulus
  - Poisson's ratio
  - Other material properties

- 68-95-99.7 rule

$$P(\mu_x - \sigma_x \le x \le \mu_x + \sigma_x) \approx 0.68$$

$$P(\mu_x - 2\sigma_x \le x \le \mu_x + 2\sigma_x) \approx 0.95$$

$$P(\mu_x - 3\sigma_x \le x \le \mu_x + 3\sigma_x) \approx 0.997$$

✓ $Z = a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n$    $Z$ is also **normal**   $\mu_z = a_0 + \sum_{i=1}^{n} a_i\mu_i$    $\sigma_z = \sqrt{\sum_{i=1}^{n}(a_i\sigma_i)^2}$

$X_i$: independent normal variables

## Standard Normal Distribution

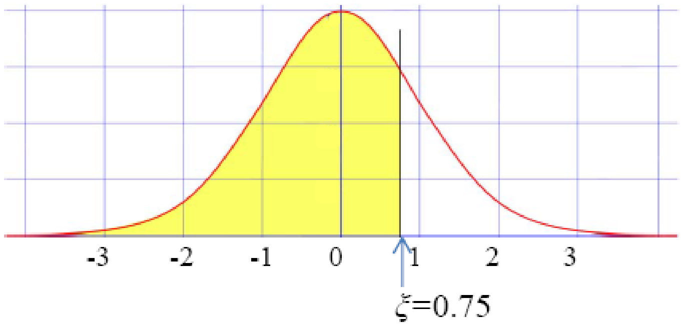
- ***Normalized normal distribution*** with zero mean and unit standard deviation, $\mathcal{N}(0,1)$
  - PDF

$$f_{\Xi}(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-\xi^2}{2}\right], \quad -\infty < \xi < \infty \qquad \xi = \frac{x-\mu_x}{\sigma_x}$$

$X$: normal variable

  - CDF

$$\Phi(\xi) = F_{\Xi}(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-\xi^2}{2}\right] d\xi$$




$$\Phi(\xi = 0.75) = 0.7734$$

Standard normal table $\Phi(\xi)$

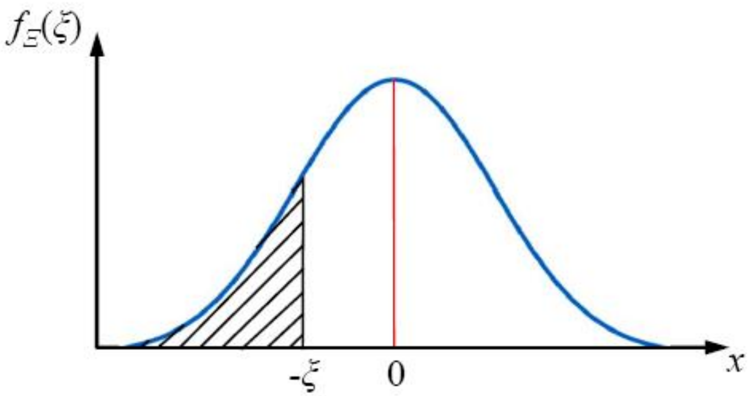
| $\xi$ | 0.03 | 0.04 | 0.05 | 0.06 | |
|---|---|---|---|---|---|
| 0.6 | 0.7357 | 0.7389 | 0.7422 | 0.7357 | . |
| 0.7 | 0.7673 | 0.7704 | **0.7734** | 0.7673 | . |
| 0.8 | 0.7967 | 0.7995 | 0.8023 | 0.7967 | . |
| . . . | | | | | |

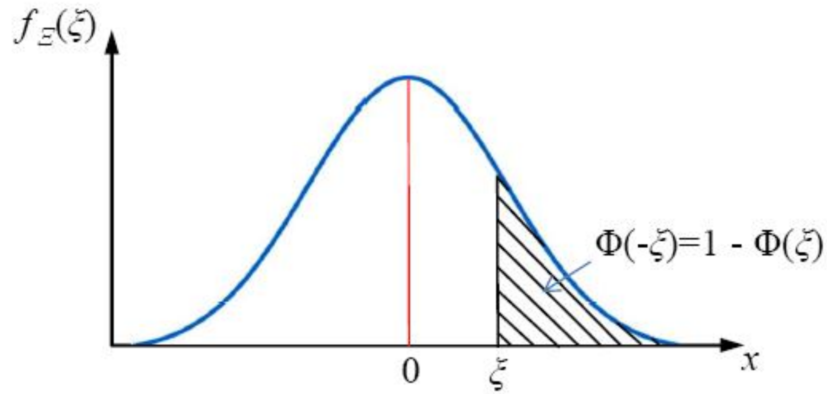
✓ If $\Phi(\xi_p) = p$ is known,

$\xi_p = \Phi^{-1}(p)$ $\quad \Phi^{-1}$: inverse of $\Phi$

ex) $\xi_p = \Phi^{-1}(0.7734) = 0.75$

# Standard Normal Distribution




$$\Phi(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-\xi^2}{2}\right] d\xi$$

- ✓ $\Phi(\xi) + \Phi(-\xi) = 1$ due to symmetry about zero

- ✓ If $\Phi(\xi) = p$,
  $\Phi(-\xi) = 1 - \Phi(\xi) = 1 - p$

  $-\xi = \Phi^{-1}(1-p)$ or
  $\xi = -\Phi^{-1}(1-p)$

  Standard normal table $\Phi(\xi)$
  is usually available for $\xi > 0$

## Lognormal distribution

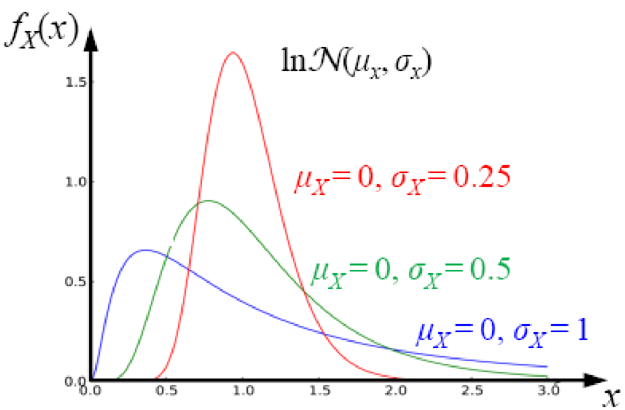
- Especially useful when ***negative values*** of a system parameter are physically ***impossible*** and ***a large range of data*** are involved
  - Examples:
    Material strength, loading variables, cycles to failure, etc.
  - $Y = \ln X$ If $X$ is a **lognormal** distribution, then $Y$ is **normally** distributed

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right], \ -\infty < y < \infty$$

$$f_X(x) = \frac{1}{x\sigma_X\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu_X}{\sigma_X}\right)^2\right], \ 0 < x < \infty$$

$$\mu_Y = \ln\mu_X - \frac{1}{2}\sigma_Y{}^2$$

$$\sigma_Y{}^2 = \ln\left[\left(\frac{\sigma_X}{\mu_X}\right)^2 + 1\right]$$




$$E[X] = \exp\left(\mu_X + \frac{1}{2}\sigma_X{}^2\right)$$

$$V[X] = \left(\exp\left(\sigma_X{}^2\right) - 1\right)\exp\left(2\mu_X + \sigma_X{}^2\right)$$

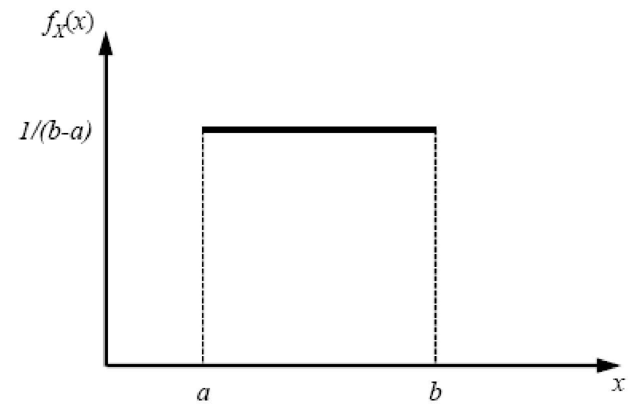
# Uniform & Gamma Distribution

## Uniform Distribution

- All outcomes are equally likely to occur

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & x < a \text{ or } x > b \end{cases} \qquad E[X] = \frac{a+b}{2}$$

$$F_X(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases} \qquad V[X] = \frac{(b-a)^2}{12}$$



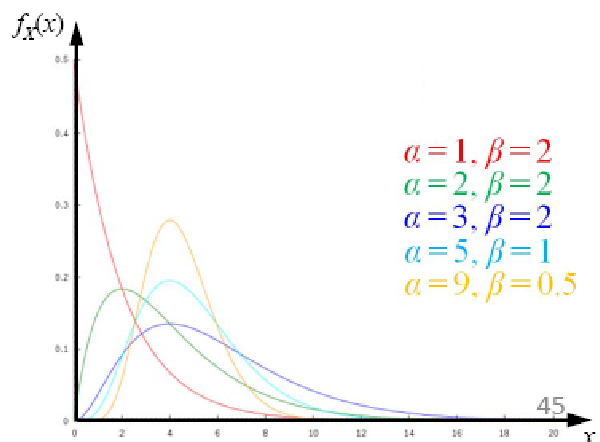

## Gamma Distribution

- Frequently used to model waiting time

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \; 0 \le x < \infty \qquad \alpha > 0, \; \beta > 0$$

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x) dx \qquad \text{Gamma function}$$

$$E[X] = \alpha\beta \qquad V[X] = \alpha\beta^2$$



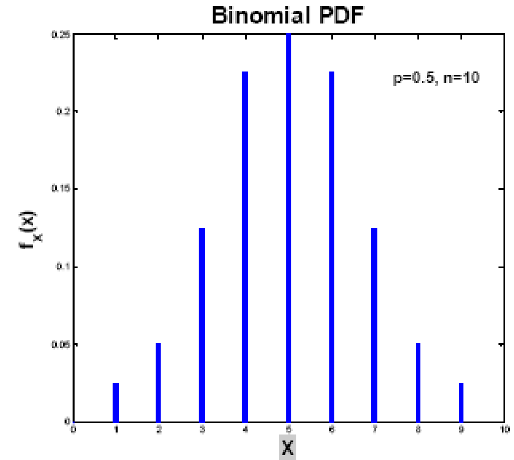
# Exponential and Binomial Distribution

## Binomial Distribution

- Generalization of Bernoulli for n independent trials
- Gives probability of success of r unfavorable outcomes in n trials

$$f_X(x) = P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

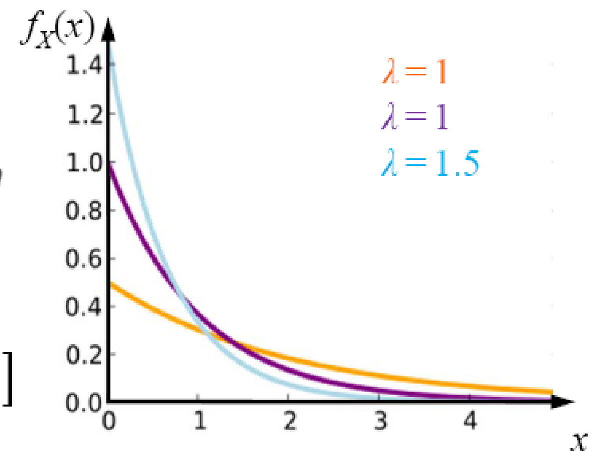



## Exponential Distribution

- Special case of ***Type III extreme value*** (or ***Weibull) distribution for $\alpha$ = 1***
- Usually describes the time between events in a ***Poisson process*** (a process in which events occur continuously and independently at a constant average rate)

$$f_X(x) = \lambda \exp[-\lambda x], x \geq 0 \qquad F_X(x) = 1 - \exp[-\lambda x]$$

$$E[X] = \frac{1}{\lambda} \qquad V[X] = \frac{1}{\lambda^2}$$



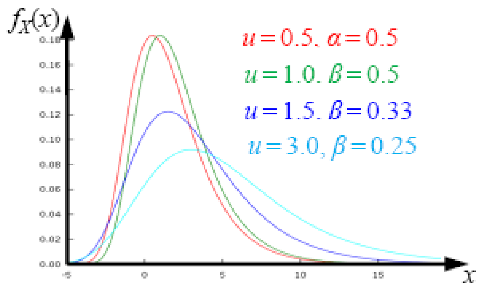
# Extreme Value Distribution



## Extreme Value Distribution

- Used as an approximation to ***model the maximum or minimum*** of long (finite) sequences of random variables
  - **Type I** extreme value (or *Gumbel*) distribution

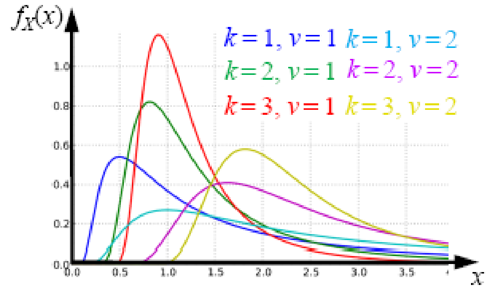
$$f_X(x) = \alpha \exp\left[-\exp\left(-\alpha\left(x-u\right)\right)\right]\exp\left[-\alpha\left(x-u\right)\right]$$




  - **Type II** extreme value (or *Fréchet*) distribution

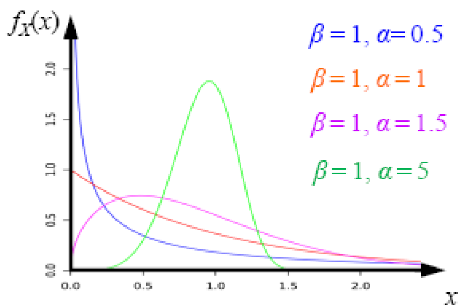
$$u = ln\ v,\ \alpha = k$$

$$f_X(x) = \frac{k}{v}\left(\frac{v}{x}\right)^{k+1}\exp\left[-\left(\frac{v}{x}\right)^{k}\right],\ 0 \le x < \infty,\ k \ge 2$$




  - **Type III** extreme value (or *Weibull*) distribution

$$f_X(x) = \frac{\alpha x^{\alpha-1}}{\beta^{\alpha}}\exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right], x \ge 0, \alpha > 0, \beta > 0$$



# Determination of Probability Distribution

- Probability Paper Method

**Table 5.1 Preparation of Data for Young's Modulus for Plotting on Probability Papers**

| $m$ | $E$ (ksi) | $m/(N+1)$ | $m$ | $E$ (ksi) | $m/(N+1)$ |
|---|---|---|---|---|---|
| 1 | 25,900 | 1/42 = 0.0238 | 21 | 29,400 | 21/42 = 0.5000 |
| 2 | 27,400 | 0.0476 | 22 | 29,400 | 0.5238 |
| 3 | 27,400 | 0.0714 | 23 | 29,500 | 0.5476 |
| 4 | 27,500 | 0.0952 | 24 | 29,600 | 0.5714 |
| 5 | 27,600 | 0.1190 | 25 | 29,600 | 0.5952 |
| 6 | 28,100 | 0.1429 | 26 | 29,900 | 0.6190 |
| 7 | 28,300 | 0.1667 | 27 | 30,200 | 0.6429 |
| 8 | 28,300 | 0.1905 | 28 | 30,200 | 0.6667 |
| 9 | 28,400 | 0.2143 | 29 | 30,200 | 0.6905 |
| 10 | 28,400 | 0.2381 | 30 | 30,300 | 0.7143 |
| 11 | 28,700 | 0.2619 | 31 | 30,500 | 0.7381 |
| 12 | 28,800 | 0.2857 | 32 | 30,500 | 0.7619 |
| 13 | 28,900 | 0.3095 | 33 | 30,600 | 0.7857 |
| 14 | 29,000 | 0.3333 | 34 | 31,100 | 0.8095 |
| 15 | 29,200 | 0.3571 | 35 | 31,200 | 0.8333 |
| 16 | 29,300 | 0.3810 | 36 | 31,300 | 0.8571 |
| 17 | 29,300 | 0.4048 | 37 | 31,300 | 0.8810 |
| 18 | 29,300 | 0.4286 | 38 | 31,300 | 0.9048 |
| 19 | 29,300 | 0.4524 | 39 | 32,000 | 0.9286 |
| 20 | 29,300 | 0.4762 | 40 | 32,700 | 0.9524 |
| | | | 41 | 33,400 | 0.9762 |

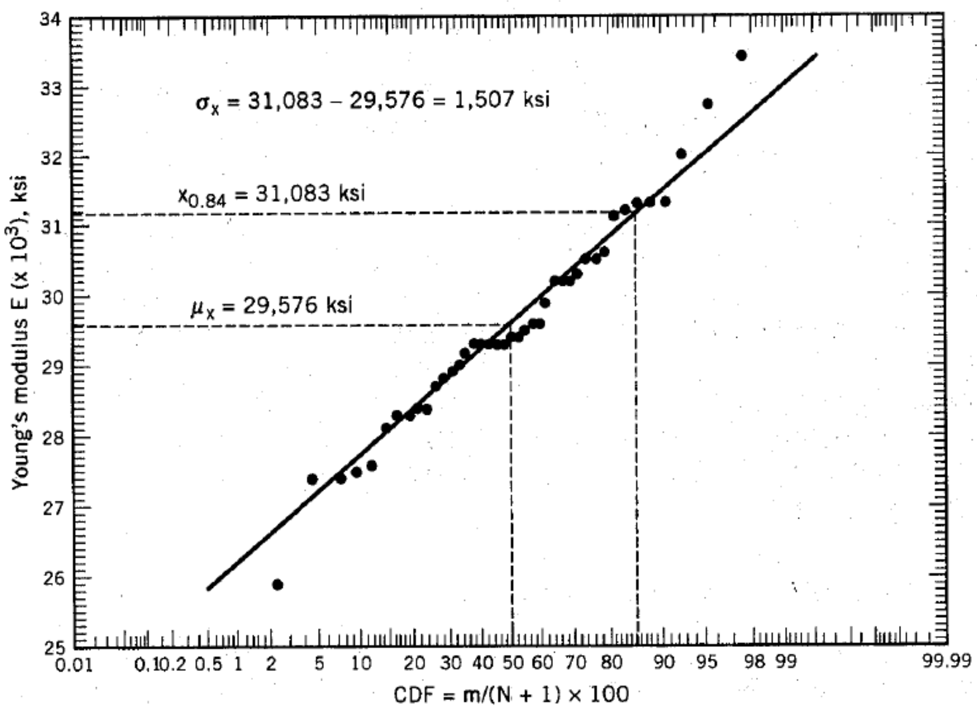


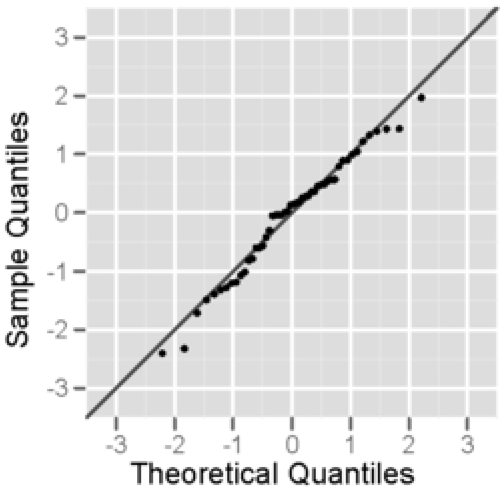

**Figure 5.1** Normal Probability Paper

# Determination of Probability Distribution

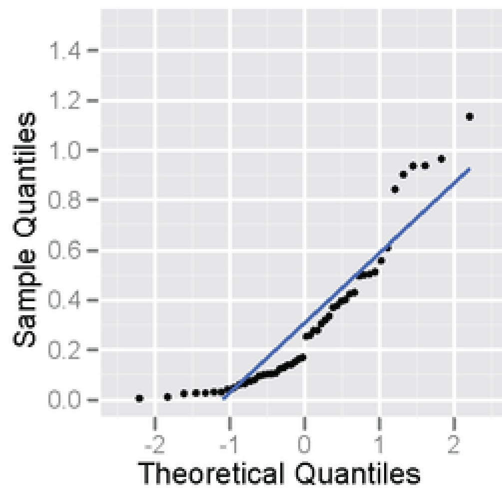


## Normal Probability Plot

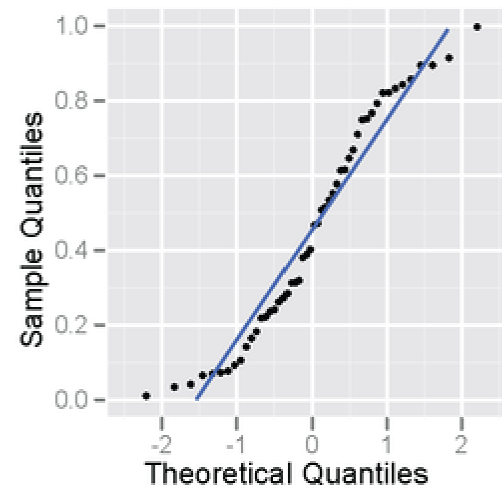
### Normal data

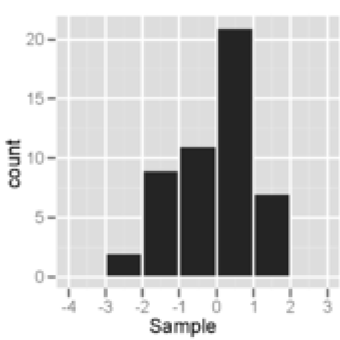


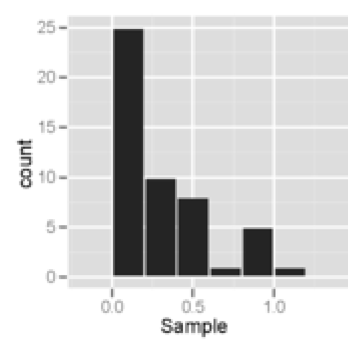

### Right-skewed

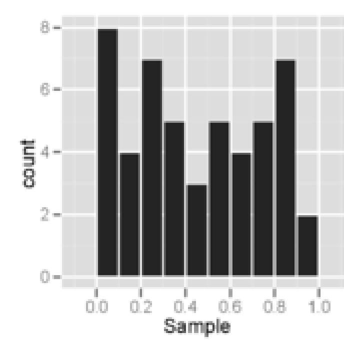



### Uniform data

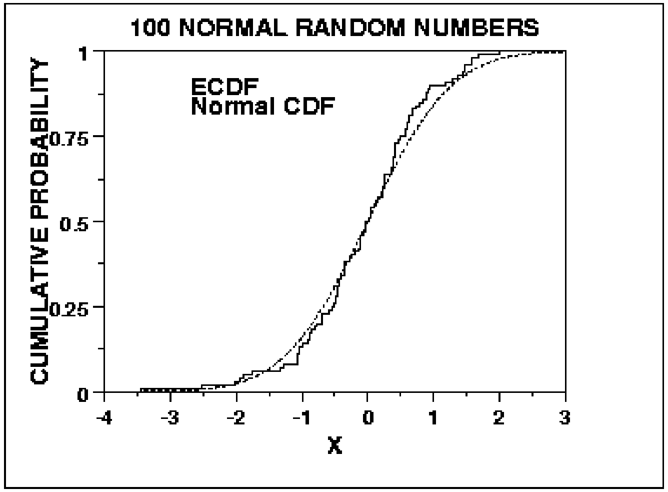


# Kolmogorov-Smirnov test




| Sample Size, N | Level of Significance, $\alpha$ | | | | |
|---|---|---|---|---|---|
| | 0.20 | 0.15 | 0.10 | 0.05 | 0.01 |
| 3 | 0.565 | 0.597 | 0.642 | 0.708 | 0.828 |
| 4 | 0.494 | 0.525 | 0.564 | 0.624 | 0.733 |
| 5 | 0.446 | 0.474 | 0.474 | 0.565 | 0.669 |
| 10 | 0.322 | 0.342 | 0.368 | 0.410 | 0.490 |
| 15 | 0.266 | 0.283 | 0.304 | 0.338 | 0.404 |
| 20 | 0.231 | 0.246 | 0.264 | 0.294 | 0.356 |
| 25 | 0.21 | 0.22 | 0.24 | 0.27 | 0.32 |
| 30 | 0.19 | 0.20 | 0.22 | 0.24 | 0.29 |
| 35 | 0.18 | 0.19 | 0.21 | 0.23 | 0.27 |
| 40 | 0.17 | 0.18 | 0.19 | 0.21 | 0.25 |
| 45 | 0.16 | 0.17 | 0.18 | 0.20 | 0.24 |
| 50 | 0.15 | 0.16 | 0.17 | 0.19 | 0.23 |
| over 50 | $\frac{1.07}{\sqrt{N}}$ | $\frac{1.14}{\sqrt{N}}$ | $\frac{1.22}{\sqrt{N}}$ | $\frac{1.36}{\sqrt{N}}$ | $\frac{1.63}{\sqrt{N}}$ |

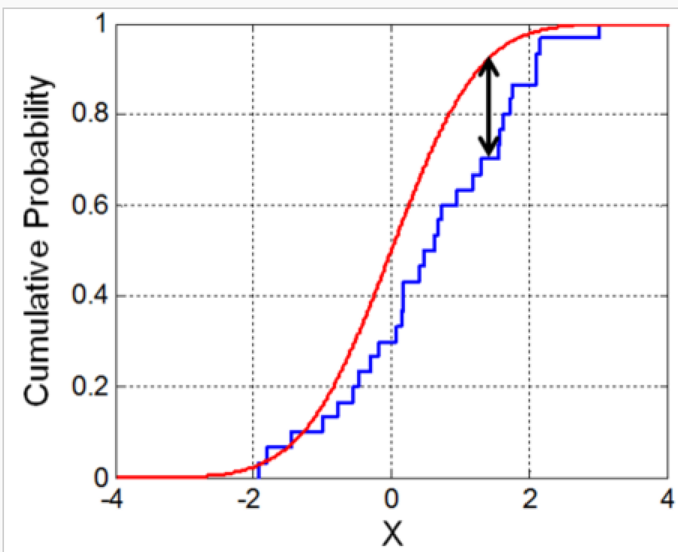



Illustration of the Kolmogorov–Smirnov statistic. Red line is CDF, blue line is an ECDF, and the black arrow is the K–S statistic.

K-S test:

$$P(D_n \le D_n^{\alpha}) = 1 - \alpha$$

# Central Limit Theorem

**Central Limit Theorem** (CLT)

- States that given certain conditions, the linear combination of a sufficiently large number of ***independent random variables*** will be ***approximately normally distributed***
- Suppose that $X_{1,}, X_2, \ldots, X_n$ are any independent random variables with the same mean $\mu$ and variance $\sigma^2$,
  then the sum and mean of all $X$'s are approximated by

$$X_1 + X_2 +, \ldots, +X_n \sim \mathcal{N}\left(n\mu, \sqrt{n\sigma^2}\right)$$

$$\frac{X_1 + X_2 +, \ldots, +X_n}{n} \sim \mathcal{N}\left(\mu, \sqrt{\frac{\sigma^2}{n}}\right)$$

Usually, CLT is valid when $n \geq 30$

Ex: When $n$ dice are rolled, the sum of $n$ dice is approximated by a normal distribution

$$X_1 + X_2 +, \ldots, +X_n \sim \mathcal{N}\left(3.5n, \sqrt{2.92n}\right)$$

$X_i$: number rolled on each die

Mean of number rolled on each die: $(1+2+3+4+5+6)/6 = 3.5$

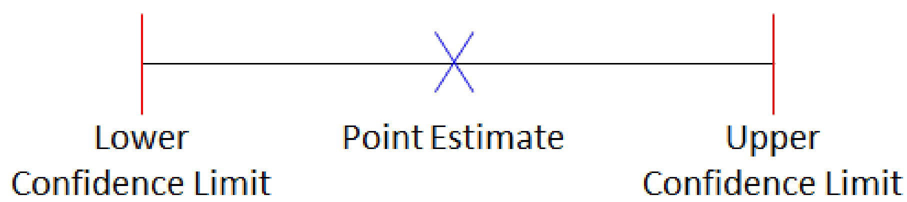
Variance of number rolled on each die: $\{(1\text{-}3.5)^2 + \cdots + (6\text{-}3.5)^2\}/6 = 2.92$

# Confidence Interval

- Based on Normal approximation
- Estimated range of values that is likely to include an unknown population parameter (e.g. mean of a population)
- Calculated from a given set of sample data
- Take into consideration variation in point estimate from sample to sample
- Stated in terms of confidence level (90%, 95%, and 99% commonly used )

e.g. 95 % confidence interval: 95 % confident that the Interval contains an unknown parameter

**Confidence Interval**




Point estimate: Best estimate of an unknown population parameter based on sample data

$$\text{Confidence interval} = \text{Point Estimate} \pm \text{Margin of Error}$$

# Confidence Interval



## Confidence interval to estimate population mean $\mu$

- Assume that
  - Population is normally distributed and its standard deviation $\sigma$ is known
  - sample size is large: $n \geq 30$ (CLT)

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

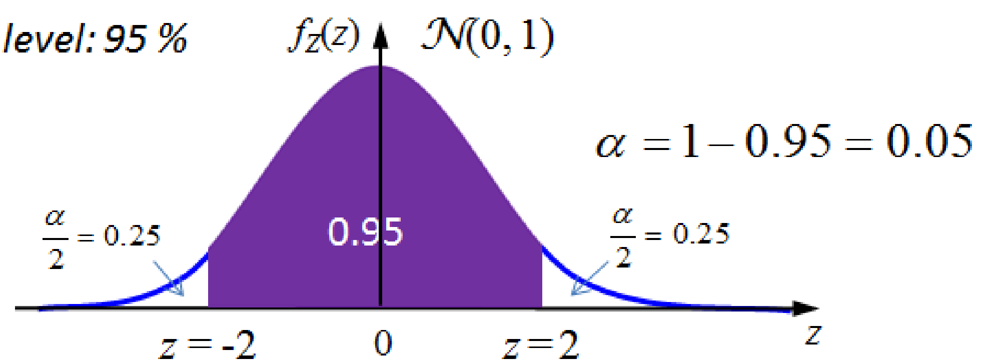
$\bar{x}$: Sample mean

$\sigma$: Population standard deviation
(if $n \geq 30$, sample standard deviation can be used instead)

$n$: Sample size

$Z_{\alpha/2}$: Standard normal distribution's critical value for probability of $\alpha/2$ in each tail

*Confidence level: 95 %*



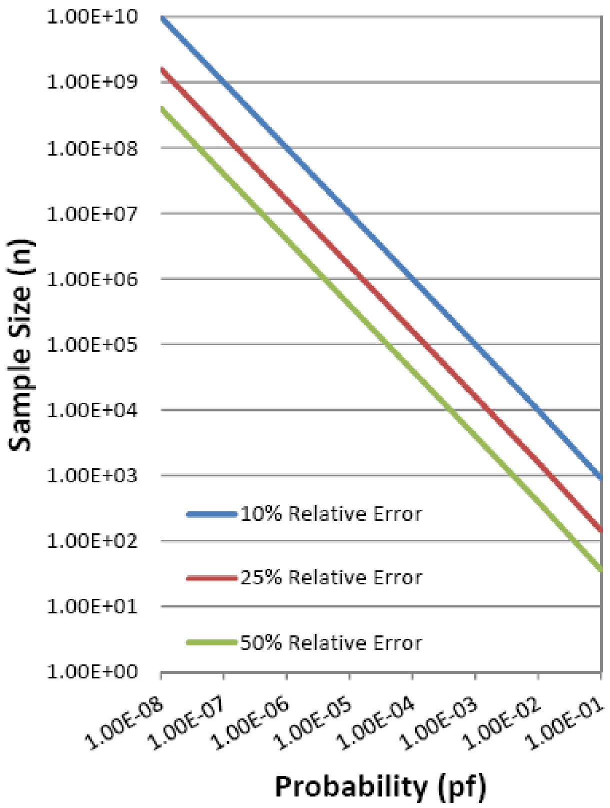
# Estimating Required Sample Size

- Required sampled size can be found mathematically
  - Depends on desired relative error
  - Depends on confidence level

$$n = \left(\frac{1 - p_f}{p_f}\right)\left(\frac{k}{Relative\ Error}\right)^2$$

$p_f = probability$
$k = confidence\ level$

| k = # of Std. Dev. | Two-Sided Confidence Level |
|---|---|
| 1 | 68% |
| 1.64 | 90% |
| 1.96 | 95% |
| 2.6 | 99% |



Maximum Likelihood Estimation

$$\mathcal{L}(\theta\,;x_1,\ldots,x_n) = f(x_1,x_2,\ldots,x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta).$$

https://en.wikipedia.org/wiki/Maximum_likelihood

Bayesian Estimation

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta}.$$

https://en.wikipedia.org/wiki/Bayes_estimator

# Fundamental Concept of Likelihood

- When $f_X(x|\theta)$ is viewed as a ***function of observable random variable*** $X$ with ***fixed parameters*** $\theta$, it is a probability distribution function of $X$

- It would be ***improper to switch likelihood and probability*** in the following two sentences
  - If I were to flip a fair coin 10 times, what is the ***probability*** of it landing heads-up every time?
  - Given that I have flipped a coin 10 times and it has landed heads-up 10 times, what is the ***likelihood*** that the coin is fair?

# Fundamental Concept of Likelihood

Likelihood can be used to ***estimate unknown statistical parameters*** given observed experimental data

$$f_X\left(x\middle|n,p\right)=\frac{n!}{x!\left(n-x\right)!}p^x\left(1-p\right)^{n-x}$$ denoted by $X\sim B\left(n,p\right)$

- When $L(\theta\,|\,x)$ is viewed as a ***function of unknown parameter*** $\theta$ with ***observed realization of random variable*** $X$ ($X$ being held at $x$), it is a likelihood function of $\theta$

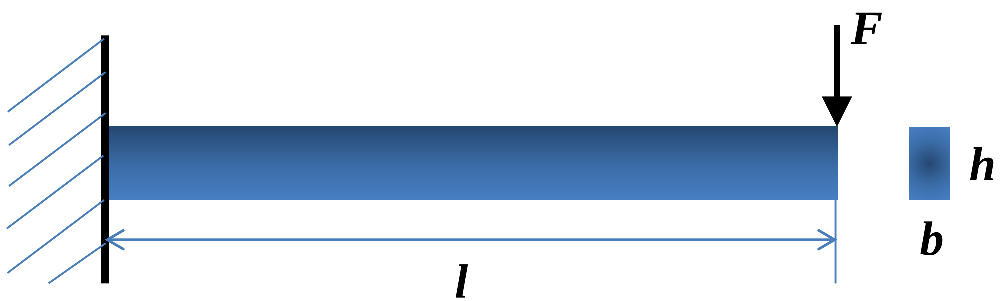
$$L\left(p\middle|X=x,n\right)=\frac{n!}{x!\left(n-x\right)!}p^x\left(1-p\right)^{n-x}$$ $X=x$ *from* $B\left(n,p\right)$

# Fundamental Concept of Likelihood

- Maximum likelihood estimator of unknown parameter $\theta$:

Defined as the ***point that maximizes likelihood function*** of $\theta$

$$\hat{\theta}_{mle} = \max L\left(\theta \middle| X = x\right)$$




$l = 1$ m

$E = 2\text{x}10^5$ MPa

$h = 0.1$ m

$b = 0.05$ m

Tip Displacement: $d = \dfrac{4Fl^3}{Ebh^3}$

Tip load $F$ is known to follow a normal distribution of standard deviation 400 N, but its mean is unknown: $F \sim \mathcal{N}(\mu, 400\text{ N})$

Tip displacement $d$ is measured from an experimental test, $d = 0.13$ cm

***What is the mean of tip load $F$?***

# Fundamental Concept of Likelihood

Likelihood can be used to ***estimate unknown statistical parameters*** given observed experimental data

- In 1922, R. A. Fisher introduced the concept of likelihood function “to ***express the state of available information*** (data) concerning the parameters of hypothetical populations”
- Likelihood function gives a ***measure of how likely*** any particular value of parameter is, given that specific data is observed
- If a certain value of parameter gives a large likelihood value, it indicates that the ***observed data favor the parameter value***
- Likelihood function (often simply likelihood) is a ***key component of Bayesian inference***

# Likelihood vs. Probability

- Likelihood is usually a synonym for probability, but a clear technical ***distinction*** exists
- It would be ***improper to switch likelihood and probability*** in the following two sentences
  - If I were to flip a fair coin 10 times, what is the ***probability*** of it landing heads-up every time?
  - Given that I have flipped a coin 10 times and it has landed heads-up 10 times, what is the ***likelihood*** that the coin is fair?

- A probability represents the ***chance of getting specific data***, not yet observed, given that a statement is true (or Statistical parameters are known)
- A likelihood represents the ***support of a given statement*** (or statistical parameters) provided by observed data

| | **Probability, *P*** | **Likelihood, *L*** |
|---|---|---|
| Parameter | Known | Unknown |
| Exptl. Data | Unknown | Known |

# Probability Distribution Function

- When $f_X(x|\theta)$ is viewed as a ***function of observable random variable*** $X$ with ***fixed parameters*** $\theta$, it is a probability distribution function of $X$

Ex: (Discrete) binomial distribution of variable $X$

**Unknown** *variable*

$$f_X(x|n,p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$ denoted by $$X \sim B(n,p)$$

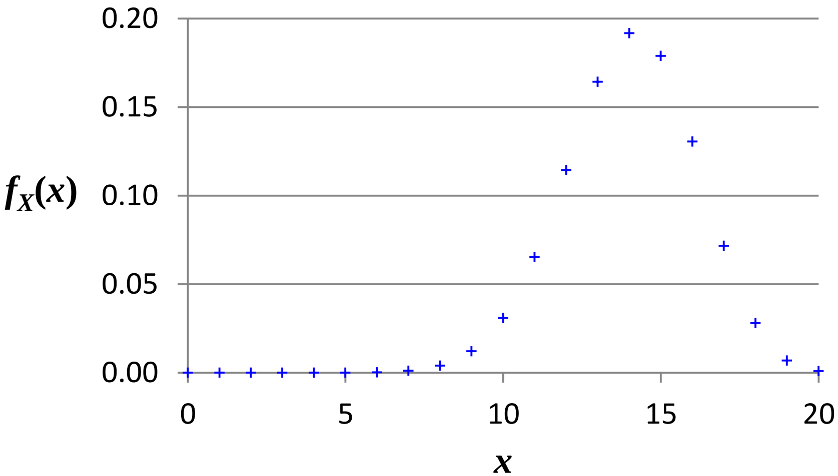
*Fixed*

$\theta$: $n, p$

$X$: Number of success
$n$: Number of trials
$p$: Success probability in each trial

Given $n$=20 and $p$=0.7

$$f_X(x|n=20, p=0.7)$$
$$= \frac{20!}{x!(20-x)!} 0.7^x (1-0.7)^{20-x}$$



# Likelihood Function

- When $L(\theta \mid x)$ is viewed as a ***function of unknown parameter*** $\theta$ with ***observed realization of random variable*** $X$ ($X$ being held at $x$), it is a likelihood function of $\theta$

Ex: Binomial likelihood function of parameter $p$

$$L\left(p \middle| X = x, n\right) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$\theta$: $p$

***Unknown*** *parameter*

$X = x$ *from* $B(n, p)$

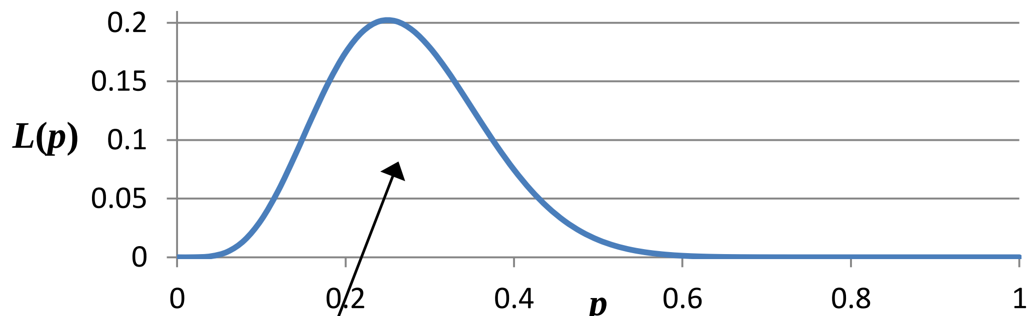
*A sample*

*Known*

(*an observed realization of* $X$)

Given $n$=20 and $X$=5

$$L\left(p \middle| n = 20, X = 5\right) = \frac{20!}{5!(20-5)!} p^5 (1-p)^{15}$$




Area under likelihood function cannot be 1 ➡ likelihood function ≠ probability distribution function

- A likelihood value is not probability (nor density) because the ***argument of likelihood function is parameter***(s) of a probability distribution function, not the random variable itself

# Maximum Likelihood Estimation (MLE)

- Solve for the set of parameter values that maximizes joint distribution of PDF

- A popular statistical inference method used to obtain the ***best (point) estimate of unknown parameter*** given observed data

- Selects the parameter value that is ***most strongly supported by observed data*** (i.e. the value most likely to have resulted in the data)

- Maximum likelihood estimator of unknown parameter $\theta$:

  Defined as the ***point that maximizes likelihood function*** of $\theta$

$$\hat{\theta}_{mle} = \max L\left(\theta \middle| X = x\right)$$

# Maximum Likelihood Estimation

- More convenient to work in terms of the natural logarithm of likelihood
  - Zero gradient of log-likelihood function is found at the same location as likelihood function thus log-likelihood function can be used to determine the maximum value of a likelihood function

In biased coin toss problem

$X = 3$
$n = 5$
$p$ = unknown

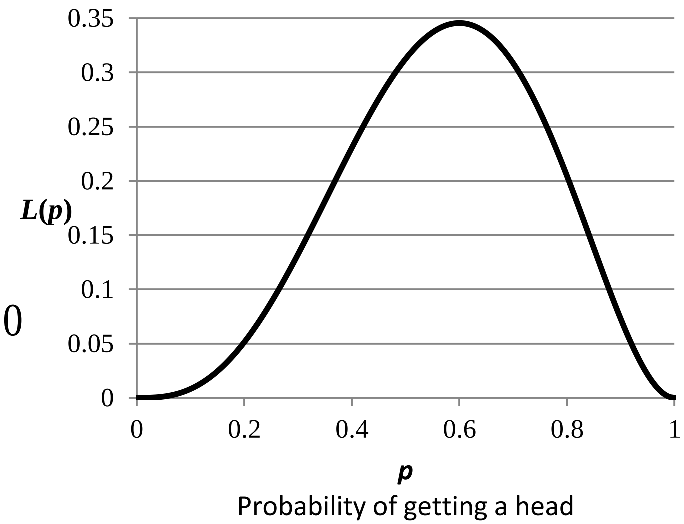
$$L\left(p \middle| X=3\right) = 10p^3(1-p)^2$$

$$\ln L\left(p \middle| X=3\right) = \ln\left(10p^3(1-p)^2\right) = \ln 10 + 3\ln p + 2\ln(1-p)$$

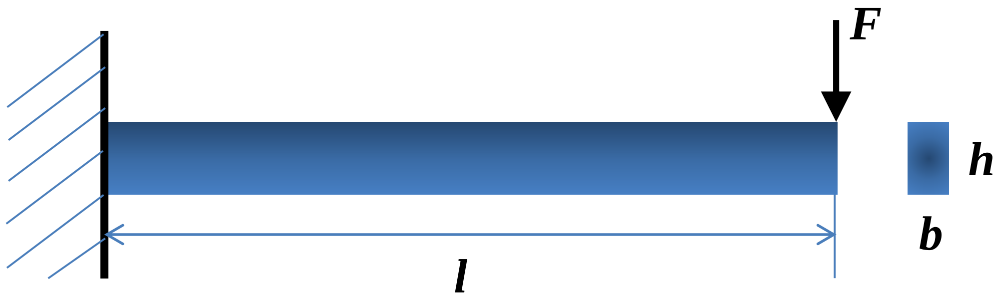
- Take derivative of the log-likelihood function and setting it to zero to find the maximum

$$\frac{d\left\{\ln L\left(p \middle| X=3\right)\right\}}{dp} = \frac{d\left\{\ln 10 + 3\ln p + 2\ln(1-p)\right\}}{dp} = \frac{3}{p} - \frac{2}{1-p} = 0$$

$$\hat{p}_{mle} = 0.6$$



# Maximum Likelihood Estimation Example




$l = 1$ m
$E = 2\text{x}10^5$ MPa
$h = 0.1$ m
$b = 0.05$ m

Tip Displacement: $d = \dfrac{4Fl^3}{Ebh^3}$

Tip load $F$ is known to follow a normal distribution of standard deviation 400 N, but its mean is unknown: $F \sim \mathcal{N}(\mu, 400\text{ N})$

Tip displacement $d$ is measured from an experimental test, $d = 0.13$ cm

***What is the mean of tip load F?***

$$d = \frac{4Fl^3}{Ebh^3}$$

$l = 1$ m, $E = 2\text{x}10^5$ MPa, $h = 0.1$ m, $b = 0.05$ m

$F \sim \mathcal{N}(\mu, 400\text{ N})$

$$d = 4\times10^{-7}F$$

# Maximum Likelihood Estimation Example

- Because tip load $F$ is a normal variable, tip displacement $d$ is also normally distributed

The mean and standard deviation of tip displacement $d$ are

$$\mu_d = 4\times10^{-7}\mu_F \qquad \sigma_d = 4\times10^{-7}\times400 = 1.6\times10^{-4}$$

$$\mu(aX) = a\mu(X)$$
$$\sigma(aX) = a\sigma(X)$$

Therefore,

$$d \sim \mathcal{N}(4\times10^{-7}\mu_F,\ 1.6\times10^{-4}\ \mathrm{N})$$

$\mu_F$ : Mean of tip load
$\mu_d$: Mean of tip displacement
$\sigma_d$: Standard deviation of tip displacement

$$f(d) = \frac{1}{\sqrt{2\pi\sigma_d^{\,2}}}\exp\left[-\frac{(d-\mu_d)^2}{2\sigma_d^{\,2}}\right] = 2.49\times10^3\times\exp\left[-1.95\times10^7\left(d-4\times10^{-7}\mu_F\right)^2\right]$$

- Given the observed data on tip displacement $d$ = 0.0013 m, likelihood function of unknown parameter $\mu_F$ is obtained by putting the $d$ value into $f(d)$

$$L\left(\mu_F \middle| d = 0.0013\right) = 2.49\times10^3\times\exp\left[-1.95\times10^7\left(0.0013-4\times10^{-7}\mu_F\right)^2\right]$$

# Maximum Likelihood Estimation Example

- Taking the logarithm of $L(\mu_F | d)$ gives

$$\ln L\left(\mu_F \middle| d = 0.0013\right) = -4.86\times10^{10}\times\left(0.0013 - 4\times10^{-7}\mu_F\right)^2$$

- To find the maximum likelihood estimator of $\mu_F$, take the derivative of the log-likelihood function with respect to $\mu_F$ and set it to zero

$$\frac{d\ln L\left(\mu_F \middle| d = 0.0013\right)}{d\mu_F} = \frac{d\left\{-4.86\times10^{10}\times\left(0.0013 - 4\times10^{-7}\mu_F\right)^2\right\}}{d\mu_F}$$

$$= 8\times4.86\times10^{3}\times\left(0.0013 - 4\times10^{-7}\mu_F\right) = 0$$

- Solving the equation for $\mu_F$ gives

$$\left(\hat{\mu}_F\right)_{mle} = 3,250\ \mathrm{N}$$

- $\left(\hat{\mu}_F\right)_{mle} = 3,250\ \mathrm{N}$ is the value that is most likely to have generated the observed tip displacement $d$ = 0.13 cm

# Summary of Likelihood

- Likelihood provides ***a measure of how likely specific values of statistical parameters*** are, given that data are observed
- In likelihood, ***random variables are fixed*** at certain values (data values are known) while some ***statistical parameters are unknown***
- Likelihood value is meaningful ***in comparison with other likelihood values***
- Maximum likelihood estimation provides ***the best point estimate*** of unknown statistical parameter given observed data
- Likelihood of each model given experimental data (***model likelihood***) is formulated based on the concept of likelihood

## Maximum Likelihood Estimation

$$\mathcal{L}(\theta\,;x_1,\ldots,x_n) = f(x_1,x_2,\ldots,x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta).$$

https://en.wikipedia.org/wiki/Maximum_likelihood

## Bayesian Estimation

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta}.$$

https://en.wikipedia.org/wiki/Bayes_estimator