HANDOUT #3 - Summaries of Population Distributions

TOPICS

1. Definition of Population/Process
2. Definition of Random Variable
3. Types of Random Variables
4. Functions which Characterize Random Variables
5. Functions Associated with Reliability/Survival Analysis
6. Interrelationships between cdf, pdf, pmf, quantile functions
7. Examples of Distributions: Discrete, Continuous, Mixtures
8. Interrelationships between Various Distributions
Definition of Population/Process/Random Variable

1. **Statistical Population** - Collection of all possible items or units possessing one or more common characteristics under specified experimental or observational conditions

2. **Process** - Repeatable series of actions that results in an observable characteristic or measurement

   Industrial and Laboratory experiments often are characterized as hypothetical populations. Why?

   Example 1:

   Example 2:

3. **Random Experiment** - Procedure or operation whose outcome is uncertain and can not be predicted in advance

4. **Sample Space** - Collection of all possible outcomes of a random experiment

5. **Random Variable (RV)** - A function from the sample space to the real line

   Assigns a unique numerical value to each element of the sample space
Example 1: Randomly select a sample of water (1 liter) from a river and record amount of PCB in the container

Example 2: Randomly select light bulb from distribution center and measure

(a) Time to failure of light bulb

(b) Amount of protective coating on bulb

(c) Determine if bulb is defective or not
6. **Types of Random Variables** - Three major classifications:

(a) **Discrete RV** - Collection of possible values of RV is at most a finite or countably infinite set

(b) **Continuous RV** - Collection of possible values of RV is one or more intervals on the real line (probability that it assumes any specific value is 0)

(c) **Discrete-Continuous Mixture** - Collection of possible values of RV is one or more intervals on the real line and a set of distinct values

Example 1: Let Y be the number of fish captured in a randomly placed net in the Gulf of Mexico divided by the length of time the net is in the water, Catch Per Unit Effort (CPUE)

Example 2: Let X be the amount spent on health insurance per member of the household by a randomly selected employee at Texas A&M University.

Example 3: Let G be the growth rate of a randomly selected plant after receiving a prescribed amount of growth stimulant

The following diagram depicts a variety of RV’s that may be defined on a single random experiment.
Population:
River in which cooling water from nuclear power plant is discharged
Assess Environmental Impact

Continuous Variables:
1. Water Temperature
2. Shell Thickness of Turtle Eggs
3. Proportion of Frogs with Mutations
4. Maximum Daily Oxygen Content
5. Survival Time of Exposed Fish
6. Weight of Eggs in Fish

Possible Models:
1. Normal
2. LogNormal
3. Beta
4. Weibull
5. Gamma
6. Normal Mixture

Discrete Variables:
1. Number of Hatched Eggs/Fish
2. Number of Undersized Fish in SRS of 50 Fish
3. Number of Skin Sores per Frog
4. Number of Days in a Month Temp > 50°C
5. Number of Turtles Examined Until First Cracked Shell is Found

Possible Models:
1. Poisson
2. Hypergeometric
3. Poisson
4. Binomial
5. Negative Binomial

Collect Data and Evaluate Fit of Models:
What was condition of river prior to discharge?
How large a sample is needed?
How accurately do we need to estimate parameters?
How powerful a test of hypotheses is required?
Predict water quality if temperature is decreased.
Is there a trend over time of water quality?
Which models fit the data best?

Report containing conclusions:
Graphical presentation of data
Assessment of accuracy of conclusions
Suggested follow up studies
Characterization/Descriptions of Populations/Processes

Let $Y$ be a R.V. associated with a Population/Process

Let $R(Y)$ be the possible values of $Y$

Three Functions Which Completely Describe $Y$:

1. The Cumulative Distribution Function (cdf) of $Y$, $F(y)$:

$$F(y) = Pr[Y \leq y], -\infty < y < \infty$$

That is, $F : (-\infty, \infty) \Rightarrow [0, 1]$

That is, $F$ maps $(-\infty, \infty)$ into $[0,1]$

2. The Probability Mass Function (pmf) for discrete r.v.’s or Probability Density Function (pdf) for continuous r.v.’s

(a) For Discrete R.V.’s:

$$p(y) = Pr[Y = y]$$

$$F(y) = \sum_{t \leq y} p(y)$$

(b) For Continuous R.V.’s, the pdf is defined as that function, $f()$ such that

$$f(y) \geq 0 \quad F(y) = \int_{-\infty}^{y} f(t)dt \quad \Rightarrow f(y) = \frac{dF(y)}{dy}$$

$$Pr[a \leq Y \leq b] = \int_{a}^{b} f(t)dt = \text{area under } f() \text{ between a and b}$$

See Graphs on following pages.
The discrete R.V. $Y$ has 5 possible values: $y_1, y_2, y_3, y_4, y_5$, which occur with probability $p_1, p_2, p_3, p_4, p_5 \Rightarrow P(Y = y_i) = p_i$. 

CDF $F(\cdot)$

Quantile Function $Q(\cdot)$

PMF $p(\cdot)$
Discrete - Continuous Mixture (point mass at $Y = y_2$)

Cumulative Distribution Function

Quantile Function

$\gamma(p)$

$\gamma(q_1) = y_1$

$\gamma(q_2) = y_2$

$\gamma(q_3) = y_3$

$\gamma(q_4) = y_4$

$\gamma(q_5) = y_5$

$p_1$, $p_2$, $p_3$, $p_4$, $1.0$
3. The **Quantile Function** of \( Y, Q(u) \):

Inverse of the cdf: \( Q(u) = F^{-1}(u) \)

\( Q : [0, 1] \to (-\infty, \infty) \)

Case 1: For a continuous, strictly increasing cdf, \( F(\cdot) \)

\[
Q(u) = y_u \quad \text{where} \quad F(y_u) = u
\]

Case 2: For a discrete or discrete-continuous mixture r.v. the inverse of the cdf may not be defined for one of the following reasons:

i. For specified \( u \in [0, 1] \) there is no real number \( y_u \) for which \( F(y_u) = u \) or

ii. For specified \( u \in [0, 1] \) there may exist many real numbers \( y_u \) for which \( F(y_u) = u \)

In either case, the inverse of \( F \) would not be valid. For example, see discrete-continuous mixture graphs on previous page, for any value \( y \) satisfying \( y_2 < y < y_3 \), \( F(y) = p_3 \).

Therefore, by the definition in Case 1,

\[
F(y_2) = p_3 \quad \text{and} \quad F(y_3) = p_3
\]

Thus, by the definition of an inverse function,

\[
Q(p_3) = y_2 \quad \text{and} \quad Q(p_3) = y_3
\]

But this violates the definition of a function (the same value in the domain is mapped to two distinct values.)

Also, for all \( u \) satisfying \( p_2 < u < p_3 \), there is no real number \( y_u \) for which \( F(y_u) = u \).

Thus, we have the following **Alternative Definition:**

For \( u \in (0, 1) \), the \( 100u \)-quantile of the r.v. \( Y \) (or cdf \( F \)) is the quantity \( y_u = Q(u) \) such that

\[
Q(u) = y_u = \inf \{ y : F(y) \geq u \}
\]

That is, \( Q(u) \) is the smallest value of \( y \) for which \( F(y) \geq u \).

In our example of a discrete-continuous mixture distribution,

1. \( Q(p_3) = y_2 \) (\( y_2 \) is the smallest value of \( y \) for which \( F(y) \geq p_3 \))
2. For all $u$ satisfying $p_2 < u < p_3$, $Q(u) = y_2$ (For $u$ satisfying $p_2 < u < p_3$, $F(y_2) = p_3 > u$ and $F(y) < u$ for all $y < y_2$).

Note, jumps in the cdf $F$ become flat regions in $Q$ and flat regions in $F$ become jumps in $Q$. Also, we have the following equivalent definition of the quantile function:

For $u \epsilon (0, 1)$, $y_u$ is the 100$u$ quantile of the r.v. $Y$ or cdf $F$ if

1. $Pr[Y \leq y_u] \geq u$ AND $Pr[Y \geq y_u] \geq 1 - u$

   or in terms of the cdf $F$

2. $F(y_u) \geq u$ AND $F(y_u^-) \leq u$
In many cases, the cdf or pdf of a r.v. \( Y \) is specified as a member of a family of distributions which are indexed by parameters:

\[
Y \quad \text{has a pdf in the family} \quad \{f(y; \theta) : \theta \in \Theta\}
\]

The following examples will illustrate this notation:

**EX1:** \( Y \) has pdf given by for \( -\infty < y < \infty \)

\[
f(y, \theta_1, \theta_2) = \frac{1}{\sqrt{2\pi \theta_2}} e^{-\frac{1}{2\theta_2}(y-\theta_1)^2} \quad \text{for} \quad \theta \in \Theta = \{(\theta_1, \theta_2) : \theta_1 \in (-\infty, \infty), \theta_2 \in (0, \infty)\}
\]

**EX2:** \( Y \) has pmf given by for \( y = 0, 1, \ldots, n \)

\[
f(y; \theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y} \quad \text{for} \quad \theta \in \Theta = [0, 1]
\]

The following are some special cases of family of distributions:

1. The parameter \( \theta \) in a family of pdf’s for the r.v. \( Y \), \( \{f_Y(y; \theta) : \theta \in \Theta\} \) is said to be a **Location Parameter** if the distribution of \( W = Y - \theta \) does not depend on \( \theta \), that is, if the pdf of \( W \) \( f_W(w) = f_Y(w + \theta; \theta) \) does not depend on \( \theta \). \( W \) is referred to as the **Standard** member of a location family if \( \theta = 0 \).

2. The parameter \( \theta \) in a family of pdf’s for the r.v. \( Y \), \( \{f_Y(y; \theta) : \theta \in \Theta\} \) is said to be a **Scale Parameter** if the distribution of \( W = Y/\theta = 1 \) does not depend on \( \theta \), that is, if the pdf of \( W \) \( f_W(w) = \theta f_Y(\theta w; \theta) \) does not depend on \( \theta \). \( W \) is referred to as the **Standard** member of a scale family if \( \theta = 1 \).

3. The parameters \( \theta_1 \) and \( \theta_2 \) in a family of pdf’s for the r.v. \( Y \), \( \{f_Y(y; \theta_1, \theta_2) : \theta \in \Theta\} \) are said to be a **Location-Scale Parameters** if the distribution of \( W = (Y - \theta_1)/\theta_2 \) does not depend on \( \theta_1 \) nor \( \theta_2 \), that is, if the pdf of \( W \) \( f_W(w) = \theta_2 f_Y(\theta_2(w - \theta_1); \theta_1, \theta_2) \) does not depend on \( \theta_1 \) nor \( \theta_2 \). \( W \) is referred to as the **Standard** member of a location-scale family if \( \theta_1 = 0 \) and \( \theta_2 = 1 \).
The following examples will illustrate these types of families:

**Example 1.** Let $Y$ have a $N(\theta, 1)$ distribution. Then $\theta$ is a location parameter and $Z$ having a $N(0, 1)$ distribution is the standard member of the family.

$$f(y, \theta, 1) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-\theta)^2} \text{ for } \theta \in (-\infty, \infty)$$

**Example 2.** Let $Y$ have an Exponential Distribution with parameter $\lambda$, that is,

$$f(y; \lambda) = \begin{cases} \frac{1}{\lambda}e^{-y/\lambda} & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases}$$

Then $\lambda$ is a scale parameter and $W$ with an exponential distribution having $\lambda = 1$ is the standard member of the family.
Example 3. Let $Y$ have a $N(2, \theta^2)$ distribution. Is $\theta$ a scale parameter in this distribution?

$$f(y, 2, \theta) = \frac{1}{\sqrt{2\pi\theta}}e^{-\frac{1}{2\theta^2}(y-2)^2} \text{ for } \theta \epsilon (0, \infty)$$

Example 4. Let $Y$ have a $N(\theta_1, \theta_2^2)$ distribution. Are $(\theta_1, \theta_2)$ location- scale parameters in this distribution?

$$f(y, \theta_1, \theta_2) = \frac{1}{\sqrt{2\pi\theta_2}}e^{-\frac{1}{2\theta_2^2}(y-\theta_1)^2} \text{ for } \theta \epsilon \Theta = \{(\theta_1, \theta_2) : \theta_1 \epsilon (-\infty, \infty), \theta_2 \epsilon (0, \infty)\}$$

There are many examples of distributions having parameters which are neither location nor scale parameters. The following tables and figures will illustrate such distributions:
# Table of Common Distributions

**From: Casella-Berger, "Statistical Inference"**

## Discrete Distributions

### Bernoulli(p)

| pmf          | \[ P(X = x|p) = p^x(1-p)^{1-x}; \quad x = 0, 1; \quad 0 \leq p \leq 1 \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = p, \quad \text{Var} X = p(1-p) \] |
| mgf          | \[ M_X(t) = (1-p) + pe^t \] |

### Binomial(n, p)

| pmf          | \[ P(X = x|n, p) = \binom{n}{x} p^x(1-p)^{n-x}; \quad x = 0, 1, 2, \ldots, n; \quad 0 \leq p \leq 1 \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = np, \quad \text{Var} X = np(1-p) \] |
| mgf          | \[ M_X(t) = [pe^t + (1-p)]^n \] |
| notes        | Related to Binomial Theorem (Theorem 3.2.2). The **multinomial** distribution (Definition 4.6.2) is a multivariate version of the binomial distribution. |

### Discrete uniform

| pmf          | \[ P(X = x|N) = \frac{1}{N}; \quad x = 1, 2, \ldots, N; \quad N = 1, 2, \ldots \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = \frac{N+1}{2}, \quad \text{Var} X = \frac{(N+1)(N-1)}{12} \] |
| mgf          | \[ M_X(t) = \frac{1}{N} \sum_{i=1}^{N} e^{it} \] |

### Geometric(p)

| pmf          | \[ P(X = x|p) = p(1-p)^{x-1}; \quad x = 1, 2, \ldots; \quad 0 \leq p \leq 1 \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = \frac{1}{p}, \quad \text{Var} X = \frac{1-p}{p^2} \] |
| mgf          | \[ M_X(t) = \frac{pe^t}{1-(1-p)e^t}, \quad t < -\log(1-p) \] |
| notes        | \[ Y = X - 1 \] is negative binomial(1,p). The distribution is *memoryless*: \[ P(X > s|X > t) = P(X > s - t) \]. |

### Hypergeometric

| pmf          | \[ P(X = x|N, M, K) = \binom{M}{x} \binom{N-M}{K-x}; \quad x = 0, 1, 2, \ldots, K; \quad M - (N - K) \leq x \leq M; \quad N, M, K \geq 0 \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = \frac{KM}{N}, \quad \text{Var} X = \frac{KM(N-M)(N-K)}{N(N-1)} \] |
| notes        | If \( K \ll M \) and \( N \), the range \( x = 0, 1, 2, \ldots, K \) will be appropriate. |

### Negative binomial(r, p)

| pmf          | \[ P(X = x|r, p) = \binom{r+x-1}{x} p^r(1-p)^x; \quad x = 0, 1, \ldots; \quad 0 \leq p \leq 1 \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = \frac{r(1-p)}{p}, \quad \text{Var} X = \frac{r(1-p)}{p^2} \] |
| mgf          | \[ M_X(t) = \left( \frac{p}{1-(1-p)e^t} \right)^r, \quad t < -\log(1-p) \] |
| notes        | An alternate form of the pmf is given by \( P(Y = y|r, p) = \binom{y-1}{r-1} p^r (1-p)^{y-r} \), \( y = r, r + 1, \ldots \). The random variable \( Y = X + r \). The negative binomial can be derived as a gamma mixture of Poissons. (See Exercise 4.34.) |

### Poisson(\( \lambda \))

| pmf          | \[ P(X = x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}; \quad x = 0, 1, \ldots; \quad 0 \leq \lambda < \infty \] |
|-------------|-------------------------------------------------|
| mean and variance | \[ \text{EX} = \lambda, \quad \text{Var} X = \lambda \] |
| mgf          | \[ M_X(t) = e^{\lambda(e^t-1)} \] |
### TABLE 3.3
**Descriptions and assumptions for the random variables of Table 3.2**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Random Variable, $X$</th>
<th>Assumptions</th>
</tr>
</thead>
</table>
| Bernoulli     | The number of successes in one Bernoulli trial                                       | 1. The probability of success is known.  
2. The trial can result in only one of two possible outcomes—success or failure.          |
| Binomial      | The number of successes in $n$ Bernoulli trials (may be approximated by the Poisson   | 1. Each trial can result in only one of two possible outcomes—a success or failure.  
2. The probability of a success, $p$, is constant from trial to trial.  
3. All trials are statistically independent.  
4. The number of trials, $n$, is a specified constant. |
|               | distribution, letting $\lambda = np$, when either $n > 20 \land p < 0.05$ or $n > 100 \land np < 10$) |                                                                                                                                           |
| Geometric     | The number of failures prior to the first success in a sequence of Bernoulli trials   | 1. Each trial can result in only one of two possible outcomes—a success or failure.  
2. The probability of a success, $p$, is constant from trial to trial.  
3. All trials are statistically independent.  
4. The sequence of trials terminates after the first success. |
| Negative      | The number of failures prior to the $r$th success in a sequence of Bernoulli trials   | 1. Each trial can result in only one of two possible outcomes—a success or failure.  
2. The probability of a success, $p$, is constant from trial to trial.  
3. All trials are statistically independent.  
4. The sequence of trials terminates after the $r$th success. |
| binomial      |                                                                                      |                                                                                                                                           |
| Hypergeometric| The number of successes in a sample of size $n$ (may be approximated by binomial       | 1. Sampling is performed without replacement from a finite set of size $N$ containing a successes.  
2. Each member of the sample can result in only one of two possible outcomes—a success or failure.  
3. The sample size, $n$, is a specified constant. |
|               | distribution when $N > 10n$)                                                        |                                                                                                                                           |
| Poisson       | The number of event occurrences during a specified period of time                     | 1. The average rate of occurrences ($\lambda > 0$) is known.  
2. Occurrences are equally likely to occur during any time interval.  
3. Occurrences are statistically independent. |

---

*From STATISTICAL ANALYSIS FOR ENGINEERS AND SCIENTISTS*  
A Computer-Based Approach  
J. WESLEY BARNES.
FIGURE 3.3

STATISTICAL ANALYSIS FOR ENGINEERS AND SCIENTISTS
A Computer-Based Approach  J. WESLEY BARNES,
### Continuous Distributions

**Beta(α, β)**

| pdf | \( f(x|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \ 0 \leq x \leq 1, \ \alpha > 0, \ \beta > 0 \) |
| mean and variance | \( \text{EX} = \frac{\alpha}{\alpha + \beta}, \ \text{Var} X = \frac{\alpha \beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \) |
| mgf | \( M_X(t) = 1 + \sum_{k=1}^{\infty} \left( \prod_{r=0}^{\infty} \frac{\alpha r + \beta r}{\alpha + \beta + r} \right) \frac{t^k}{k!} \) |
| notes | The constant in the beta pdf can be defined in terms of gamma functions, \( B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \). Equation (3.2.18) gives a general expression for the moments. |

**Cauchy(θ, σ)**

| pdf | \( f(x|\theta, \sigma) = \frac{1}{\pi} \frac{1}{1 + (\frac{x - \theta}{\sigma})^2}, \ -\infty < x < \infty; \ -\infty < \theta < \infty, \ \sigma > 0 \) |
| mean and variance | do not exist |
| mgf | does not exist |
| notes | Special case of Student’s t, when degrees of freedom is 1. Also, if X and Y are independent n(0, 1), X/Y is Cauchy. |

**Chi squared(ν)**

| pdf | \( f(x|\nu) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}} x^{\nu/2-1}e^{-x/2}, \ 0 \leq x < \infty; \ \nu = 1, 2, \ldots \) |
| mean and variance | \( \text{EX} = \nu, \ \text{Var} X = 2\nu \) |
| mgf | \( M_X(t) = \left( \frac{1}{1 - \frac{t}{2}} \right)^{\nu/2}, \ t < \frac{1}{2} \) |
| notes | Special case of the gamma distribution. |

**Double exponential(μ, σ)**

| pdf | \( f(x|\mu, \sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \ -\infty < x < \infty; \ -\infty < \mu < \infty, \ \sigma > 0 \) |
| mean and variance | \( \text{EX} = \mu, \ \text{Var} X = 2\sigma^2 \) |
| mgf | \( M_X(t) = \left( \frac{1}{1 - \frac{t}{\sigma}} \right)^{\alpha}, \ |t| < \frac{1}{\sigma} \) |
| notes | Also known as the Laplace distribution. |

**Exponential(β)**

| pdf | \( f(x|\beta) = \frac{1}{\beta} e^{-x/\beta}, \ 0 \leq x < \infty, \ \beta > 0 \) |
| mean and variance | \( \text{EX} = \beta, \ \text{Var} X = \beta^2 \) |
| mgf | \( M_X(t) = \frac{1}{1 - \frac{t}{\beta}}, \ t < \frac{1}{\beta} \) |
| notes | Special case of the gamma distribution. Has the memoryless property. Has many special cases: \( Y = X^{1/\gamma} \) is Weibull, \( Y = \sqrt{2X/\beta} \) is Rayleigh, \( Y = \alpha - \gamma \log(X/\beta) \) is Gumbel. |

**F**

| pdf | \( f(x|\nu_1, \nu_2) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right) \left(\frac{\nu_2}{\nu_1}\right)^{\frac{\nu_1}{2}} \left(1 + \frac{\nu_2}{\nu_1 x}\right)^{-\frac{\nu_1 + \nu_2}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}; \ 0 \leq x < \infty, \ \nu_1, \nu_2 > 1, \ldots \) |
| mean and variance | \( \text{EX} = \frac{\nu_2}{\nu_1 - 2}, \ \nu_2 > 2, \) |
| Var X = \( \frac{\nu_2}{\nu_1 - 2} \left(1 + 2\frac{\nu_2}{\nu_1 - 4}\right), \ \nu_2 > 4 \) |
| moments | \( \text{EX}^n = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)\Gamma\left(\frac{\nu_1 + 2n}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2 + 2n}{2}\right)} \left(\frac{\nu_2}{\nu_1}\right)^n, \ n \leq \frac{\nu_2}{2} \) |
| notes | Related to chi squared \( (F_{\nu_1, \nu_2}) = \left(\frac{\chi_1^2}{\nu_1} \right) / \left(\frac{\chi_2^2}{\nu_2} \right) \), where the \( \chi^2 \)s are independent and \( F \) \( (F_{1, \nu}) = \frac{e^2}{\nu} \). |

**Gamma(α, β)**

| pdf | \( f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}, \ 0 \leq x < \infty, \ \alpha, \beta > 0 \) |
| mean and variance | \( \text{EX} = \alpha\beta, \ \text{Var} X = \alpha\beta^2 \) |
| mgf | \( M_X(t) = \left( \frac{1}{1 - \frac{t}{\beta}} \right)^{\alpha}, \ t < \frac{1}{\beta} \) |
| notes | Some special cases are exponential \( (\alpha = 1) \) and chi squared \( (\alpha = p/2, \beta = 2) \). If \( \alpha = \frac{3}{2}, \ Y = \sqrt{X/\beta} \) is Maxwell. \( Y = 1/X \) has the inverted gamma distribution. Can also be related to the Poisson (Example 3.2.1). |
### Logistic(\(\mu, \beta\))

**pdf**  
\[ f(x|\mu, \beta) = \frac{1}{\beta} \frac{e^{-(x-\mu)/\beta}}{1 + e^{-(x-\mu)/\beta}} \], \(-\infty < x < \infty, \ -\infty < \mu < \infty, \ \beta > 0\)

**mean and variance**  
\[ \text{EX} = \mu, \quad \text{Var} X = \frac{\pi^2 \beta^2}{3} \]

**mgf**  
\[ M_X(t) = e^{\mu t} \Gamma(1 - \beta t) \Gamma(1 + \beta t), \quad |t| < \frac{1}{\beta} \]

**notes**  
The cdf is given by \( F(x|\mu, \beta) = \frac{1}{1 + e^{-(x-\mu)/\beta}} \).

### Lognormal(\(\mu, \sigma^2\))

**pdf**  
\[ f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} \frac{e^{-(\ln x - \mu)^2/(2\sigma^2)}}{x}, \quad 0 \leq x < \infty, \ -\infty < \mu < \infty, \ \sigma > 0 \]

**mean and variance**  
\[ \text{EX} = e^{\mu + (\sigma^2/2)}, \quad \text{Var} X = e^{2}\mu + \sigma^2 - e^{2}\mu + \sigma^2 \]

**moments**  
\( \text{(mgf does not exist)} \)

**EX}^n = e^{n\mu + n\sigma^2/2} \]

**notes**  
Example 2.3.5 gives another distribution with the same moments.

### Normal(\(\mu, \sigma^2\))

**pdf**  
\[ f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty, \ -\infty < \mu < \infty, \ \sigma > 0 \]

**mean and variance**  
\[ \text{EX} = \mu, \quad \text{Var} X = \sigma^2 \]

**mgf**  
\[ M_X(t) = e^{t\mu + t^2/2} \]

**notes**  
Sometimes called the Gaussian distribution.

### Pareto(\(\alpha, \beta\))

**pdf**  
\[ f(x|\alpha, \beta) = \frac{\beta \alpha^\beta}{x^{\beta+1}}, \quad a < x < \infty, \quad \alpha > 0, \quad \beta > 0 \]

**mean and variance**  
\[ \text{EX} = \frac{\beta \alpha}{\beta-1}, \quad \beta > 1, \quad \text{Var} X = \frac{\beta \alpha^2}{(\beta-1)^2(\beta-2)}, \quad \beta > 2 \]

**mgf**  
do not exist

### Weibull(\(\gamma, \beta\))

**pdf**  
\[ f(x|\gamma, \beta) = \frac{\beta x^{\gamma-1} e^{-x^{\beta}}}{\Gamma(\gamma)}, \quad 0 \leq x < \infty, \quad \gamma > 0, \quad \beta > 0 \]

**mean and variance**  
\[ \text{EX} = \beta^{1/\gamma} \Gamma \left(1 + \frac{1}{\gamma}\right), \quad \text{Var} X = \beta^{2/\gamma} \left[ \Gamma \left(1 + \frac{2}{\gamma}\right) - \Gamma^2 \left(1 + \frac{1}{\gamma}\right) \right] \]

**moments**  
\[ \text{EX} = \beta^{\gamma/\gamma} \Gamma \left(1 + \frac{\gamma}{\beta}\right) \]

**notes**  
The mgf exists only for \( \gamma \geq 1 \). Its form is not very useful. A special case is exponential (\( \gamma = 1 \)).

---

*From: Casella- Berger, "Statistical Inference*
**Gamma**

\[ f(x; \alpha, \beta) \]

- \( \alpha = 2, \beta = 1/3 \)
- \( \alpha = 1, \beta = 1 \)
- \( \alpha = 2, \beta = 2 \)
- \( \alpha = 2, \beta = 1 \)

*Gamma density functions*

\[ f(x; \alpha) \]

- \( \alpha = 1 \)
- \( \alpha = .5 \)
- \( \alpha = 2 \)
- \( \alpha = 5 \)

*Standard gamma density functions*

**Lognormal**

\[ f(x; \mu, \sigma) \]

- \( \mu = 1, \sigma = 1 \)
- \( \mu = 3, \sigma = \sqrt{3} \)
- \( \mu = 3, \sigma = 1 \)

*Graphs of the lognormal p.d.f.*

\( \beta = 1 \)

**Weibull**

\[ f(x; \alpha, \beta) \]

- \( \alpha = 2 \)
- \( \beta = .5 \)
- \( \beta = 10 \)

*Weibull density functions*

\[ f(x; \alpha) \]

- \( \alpha = 1 \)
- \( \beta = 10 \)
- \( \beta = 1 \)

*Exponential density functions*

\[ \lambda = 1/\beta \]

**Exponential**

**Beta**

\[ f(x; \alpha, \beta) \]

- \( \alpha = .5 \)
- \( \beta = .5 \)
- \( \beta = 2 \)

*Graphs of Weibull p.d.f.'s*

\[ \beta \equiv (\beta)^{\frac{1}{\alpha}}, \quad \alpha = \gamma \]

*Graphs of standard beta p.d.f.'s*
**FIGURE 4-5**
Representative $\chi^2$ distribution plots with low degrees of freedom.

**FIGURE 4-6**
The relationship between the standard normal distribution and the $t$ distribution.

**FIGURE 4-7**
Plots of selected $F$ distributions.
Relationships among common distributions. Solid lines represent transformations and special cases, dashed lines represent limits. Adapted from Leemis (1986).

From Casella - Berger, "Statistical Inference"
Mixture Distribution

In a number of situations neither a discrete nor a continuous distribution will adequately model the population/process. The population/process is a combination of the realizations from several discrete and/or continuous distributions. We model such situations using Mixture Distributions. In other situations, we may have several populations/processes producing the observed data.

The following examples will these situations.

Example 1: A central warehouse receives the output from several production facilities, e.g., Firestone Tires. The tires are inspected and $D_{ij}$ is the deviation from the specified adhesive strength of Tire $j$ from Facility $i$. Suppose $X_{ij}$ has a $N(\mu_i, \sigma_i^2)$ distribution, that is, the distribution of $X$ may be differ from facility to facility. Let $p_{ij}$ be the proportion tires in the warehouse from Facility $i$ with $\sum_i p_i = 1$. Let $X$ be the measurement obtained from a randomly selected tire in the central warehouse. What is the distribution of $X$?

In general, if the population of interest is a combination of the values from $k$ distinct populations, where Population $i$ has pdf $f_i$ and cdf $F_i$, then the Mixture Population has cdf and pdf given by

$$F(x) = \sum_{i=1}^{k} p_i F_i(x) \quad f(x) = \sum_{i=1}^{k} p_i f_i(x).$$

The graphs on the following pages illustrate mixtures of two normal populations.
Simulate Observation from Strictly Increasing Continuous cdf $F$

Let $Y$ have a strictly increasing continuous cdf $F_Y$. Then, the quantile function of $Y$, $Q_Y(u) = F_Y^{-1}(u)$ is a well defined function. Let $U$ have a Uniform on $(0,1)$ distribution and define the r.v. $W = Q_Y(U)$. Claim $W$ has cdf $F_W(w) = F_Y(w)$ for all $w$. That is, $W$ is a realization from the distribution of $Y$. Thus, we only need a method for generating observations from the Uniform on $(0,1)$ distribution in order to obtain realizations from any continuous distribution.

Example: Generate observations from an Exponential distribution with $\lambda = 4$. 
Simulation from Discrete cdf F

Let $D$ have a discrete distribution with cdf $F$ and pmf

$$f(d_i) = p_i \quad \text{for} \quad d_1 < d_2 < \cdots < d_k$$

$$F(d) = \sum_{d_i \leq d} p_i$$

Generate an observation from a Uniform $(0,1)$ distribution, $U$. To obtain an observation on $D$, let

$$D = \begin{cases} 
  d_1 & \text{if } U < F(d_1) \\
  d_2 & \text{if } F(d_1) < U \leq F(d_2) \\
  \quad \quad \vdots \\
  d_i & \text{if } F(d_{i-1}) < U \leq F(d_i) \\
  \quad \quad \vdots \\
  d_k & \text{if } U \leq F(d_k)
\end{cases}$$

Prove that the above method results in $D$ having cdf $F$.

Example: Generate observations from a geometric distribution with $p = .2$:

$$f(i) = p(1 - p)^{i-1} = (.2)(.8)^{i-1} \quad \text{for } i = 1, 2, 3, \ldots$$

$$F(i) = \sum_{k=1}^{i} f(k) = 1 - (1 - p)^i$$

Suppose we observe $U = .63$ then the corresponding realization from a Geometric distribution with $p = .2$ would be that integer $C$ such that $F(C - 1) < .63 \leq F(C)$. Determine $C$:

$$F(1) = .2, F(2) = .36, F(3) = .488, F(4) = .590, F(5) = .672 \quad \Rightarrow C = 5$$

The following several pages provide SAS and r code for generating random samples from specified distributions and drawing graphs of pdfs and cdfs.
S-PLUS has a systematic naming of functions related to distributions. The functions related to the same distribution have the same name except for the first letter. The first letter indicates what type of function it is. Table 6.6 explains the system.

Table 6.6. Categorization of distribution-related functions in S-PLUS

<table>
<thead>
<tr>
<th>Type1</th>
<th>Function Type</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>Distribution function</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>Probability function or cumulated density function</td>
<td></td>
</tr>
<tr>
<td>q</td>
<td>Quantile function, the inverse of the probability function</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>Random number generation</td>
<td></td>
</tr>
</tbody>
</table>

1 The significant letter for identifying the functionality of the distribution function is either d, p, q, or r. See also Table 6.7.

All these functions take numbers and vectors as input arguments. For a vector, the corresponding value for each element is computed. Then, to identify a specific function, we need to know the abbreviations for the distributions available to us, which are listed in Table 6.7.

From: "Modern Applied Statistics with S" by W.N. Venables and B.D. Ripley

Table 6.7. List of distribution-related functions in S-PLUS

<table>
<thead>
<tr>
<th>Distribution</th>
<th>S-PLUS Abbreviation</th>
<th>Parameters2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Continuous Distributions</td>
<td></td>
</tr>
<tr>
<td>Beta</td>
<td>beta</td>
<td>shape1</td>
</tr>
<tr>
<td>Cauchy</td>
<td>cauchy</td>
<td>location=0</td>
</tr>
<tr>
<td>Chi-square</td>
<td>chisq</td>
<td>df</td>
</tr>
<tr>
<td>Exponential</td>
<td>exp</td>
<td>rate=1</td>
</tr>
<tr>
<td>F</td>
<td>f</td>
<td>df1</td>
</tr>
<tr>
<td>Gamma</td>
<td>gamma</td>
<td>shape</td>
</tr>
<tr>
<td>Logistic</td>
<td>logis</td>
<td>location=0</td>
</tr>
<tr>
<td>Lognormal</td>
<td>lnorm</td>
<td>meanlog=0</td>
</tr>
<tr>
<td>Normal</td>
<td>norm</td>
<td>mean=0</td>
</tr>
<tr>
<td>Stable</td>
<td>stab</td>
<td>index</td>
</tr>
<tr>
<td>Student's t</td>
<td>t</td>
<td>df</td>
</tr>
<tr>
<td>Uniform</td>
<td>unif</td>
<td>min=0</td>
</tr>
<tr>
<td>Weibull</td>
<td>weibull</td>
<td>shape</td>
</tr>
<tr>
<td></td>
<td>Discrete Distributions</td>
<td></td>
</tr>
<tr>
<td>Binomial</td>
<td>binom</td>
<td>size</td>
</tr>
<tr>
<td>Geometric</td>
<td>geom</td>
<td>prob</td>
</tr>
<tr>
<td>Hypergeometric</td>
<td>hyper</td>
<td>m, n, k</td>
</tr>
<tr>
<td>Negative Binomial</td>
<td>nbinom</td>
<td>size</td>
</tr>
<tr>
<td>Poisson</td>
<td>pois</td>
<td>lambda</td>
</tr>
<tr>
<td>Discrete Uniform</td>
<td>sample</td>
<td>x, size=n, replace=T</td>
</tr>
<tr>
<td>Wilcoxon Rank Sum</td>
<td>wilcox</td>
<td>m</td>
</tr>
</tbody>
</table>

1 The characteristic letter, one of d, p, q, r, plus the listed abbreviation, gives the name of the S-PLUS function.

2 If parameters are preset with a default value, like in mean=0, they do not need to be specified unless another parameter value than the default is desired.

Tables 6.6 and 6.7 list the function types available. To determine the function's name, combine the characteristic letter from Table 6.6 with the abbreviation in Table 6.7. For example, to generate random numbers from the Normal distribution, the function to use is rnorm ("r" + "norm").

There are four functions related to the Normal distribution,

dnorm for calculating the value of the distribution function,
pnorm for calculating the value of the probability function,
qnorn for calculating the inverse probability function, and
rnorm for generating random numbers from the Normal distribution.
The following SAS program will generate 10 observations from a $N(0,1)$ distribution and a single observation from a Uniform on $(0,1)$ distribution:

```
DATA; DO I=1 TO 10;
  X=RANNOR(0);OUTPUT;END;
PROC PRINT;RUN;
DATA;
U=RANUNI(0);RUN;
PROC PRINT;RUN
```

There are many other distributions from which we can generate data:

- `RANPOI(0,L)`: Poisson with parameter $L$
- `RANTRI(0,H)`: Triangular with parameter $H$
- `RANBIN(0,N,P)`: Binomial with parameters $N$ and $P$
- `RANCAU(0)`: Cauchy with $\text{LOC}=0$ and $\text{SCALE}=1$
- `RANEXP(0)`: Exponential with $\text{SCALE}=1$
- `RANGAMMA(0,A)`: Gamma with $\text{SHAPE}=A$, $\text{SCALE}=1$

The following Splus program will generate 10 observations from a $N(0,1)$ distribution and a single observation from a Uniform on $(0,1)$ distribution:

```
y <- rnorm(10,0,1)
u <- runif(1,0,1)
```

Data from other distributions can be generated using the functions given on the previous handout.
In survival analysis and reliability theory, we are interested in the time to the occurrence of an event: death, failure of a machine, cancer-free examination. Let $T$ be the time at which the event occurs, with cdf $F$ and pdf $f$, then three functions related to the r.v. $T$ are

1. Survival Function is the probability that the event occurs at or after time $t$:
   \[ S(t) = Pr[T \geq t] = 1 - F(t^-) \]
   (probability device works at least $t$ units of time).

2. Hazard Function (Failure Rate or Intensity Function) is the risk of failure of a device at time $t$ given device is working at time $t$:
   \[ h(t) = \frac{f(t)}{S(t)} \Rightarrow \Delta t h(t) \approx Pr[T \leq t + \Delta t | T \geq t] \]
   $h(t)$ is generally reported as the number of failures per unit of time. It specifies the instantaneous rate of failure at time $t$ given that the device is working at time $t$.

3. Cumulative Hazard Function:
   \[ H(t) = \int_0^t h(\tau)d\tau \]
   Accumulated instantaneous risk at time $t$.

The table and graphs on the next pages will further illustrate these concepts.
**LIFETIME DISTRIBUTION REPRESENTATION RELATIONSHIPS**

<table>
<thead>
<tr>
<th></th>
<th>PDF $f(t)$</th>
<th>SURVIVAL $S(t)$</th>
<th>HAZARD RATE $h(t)$</th>
<th>CUM. HAZARD $H(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(t)$</td>
<td>$\cdot$</td>
<td>$S(t) = \int_t^\infty f(\tau)d\tau$</td>
<td>$h(t) = \frac{f(t)}{S(t)}$</td>
<td>$H(t) = \int_t^\infty f(\tau)d\tau$</td>
</tr>
<tr>
<td>$S(t)$</td>
<td>$S(t) = -S'(t)$</td>
<td>$\cdot$</td>
<td>$h(t) = \frac{-S'(t)}{S(t)}$</td>
<td>$H(t) = -\log S(t)$</td>
</tr>
<tr>
<td>$h(t)$</td>
<td>$h(t) = h(t)e^{-H(t)}$</td>
<td>$S(t) = e^{-H(t)}$</td>
<td>$\cdot$</td>
<td>$H(t) = \int_0^t h(\tau)d\tau$</td>
</tr>
<tr>
<td>$H(t)$</td>
<td>$f(t) = H'(t)e^{-H(t)}$</td>
<td>$S(t) = e^{-H(t)}$</td>
<td>$h(t) = H'(t)$</td>
<td>$\cdot$</td>
</tr>
</tbody>
</table>

The above are for a continuous distribution ($F$ and $S(t)$)

$F(t) = 1 - S(t) = \int_{-\infty}^t f(\tau)d\tau$ for all $-\infty < t < \infty$

**Note:** $h(t)$—Hazard Rate—is often denoted as $\lambda(t)$ and called Failure Rate

For discrete distributions:

1. **PMF**
   
   $p(t) = Pr\{T = t\}$ with $p(t) > 0$, $\sum_{t} p(t) = 1$

2. **CMF**
   
   $F(t) = Pr\{T \leq t\} = \sum_{\tau \leq t} p(\tau)$ for $-\infty < t < \infty$

3. **Survival**
   
   $S(t) = Pr\{T \geq t\} = \sum_{\tau \geq t} p(\tau)$ for $-\infty < t < \infty$
**Exponential Dist.**
\[ \lambda(x) = \frac{1}{\beta} \quad \text{for } x > 0 \]
\[ f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \quad \text{for } x > 0 \]

**Ex** \( \beta = 5 \)
\[ \lambda(x) = 0.2 \]

**Weibull Dist.**
\[ \lambda(x) = \gamma x^{\gamma - 1}/\beta \quad \text{for } x > 0 \]
\[ f(x) = \frac{x}{\beta} x^{\gamma - 1} e^{-\frac{x}{\beta}} \quad \text{for } x > 0 \]

**Ex** \( \gamma = 1.8 \), \( \beta = 28.26 \)
\[ \lambda(x) = 1.8 \times 18 / 28.26 = 0.637 x^{1.8} \]

**Gompertz Dist.**
\[ \lambda(x) = c e^{bx} \quad \text{for } x > 0 \]
\[ f(x) = c e^{bx - c} \left\{ e^{bx} - 1 \right\} \quad \text{for } x > 0 \]

**Ex** \( b = \ln(1.2) \) \( c = 0.087 \)
\[ \lambda(x) = 0.087 e^{1.823 x} \quad \text{for } x > 0 \]

**FIGURE 3.3-2** Plot of the failure rate \( \lambda(x) \) and the p.d.f. \( f(x) \) of the following three distributions: (a) exponential with \( \beta = 5 \); (b) Weibull with \( \beta = 6.4 \) and \( \gamma = 1.8 \); (c) Gompertz with \( b = \ln(1.2) \) and \( c = 0.087 \).