Duxbury Thomson Learning

Making Hard Decision

Third Edition

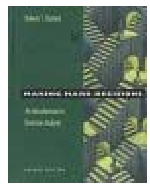
CHAPTER

9

Theoretical Probability Models

A. J. Clark School of Engineering •Department of Civil and Environmental Engineering

FALL 2003



By
Dr . Ibrahim. Assakkaf

ENCE 627 – Decision Analysis for Engineering

Department of Civil and Environmental Engineering

University of Maryland, College Park

Other Continuous Probability Distributions

- These distribution are classified as
  - Distribution used in statistical analyses
    - Student-*t*, or *t* Distribution
    - *F* Distribution
    - *Chi-square* (χ2) Distribution
  - Extreme Value Distribution
    - Type I
    - Type II
    - Type III

Other Continuous Probability Distributions

- Student-*t*, or *t* Distribution
  - The student-*t* or *t* distribution is a symmetric, bell-shaped distribution with the following density function

$$f_T(t) = \frac{\Gamma[(k+1)k]}{(\pi k)^{0.5}\Gamma(k/2)\left[1+(t^2/k)\right]^{0.5(k+1)}} \qquad -\infty < t < +\infty$$

where k is a parameter, and the gamma function $\Gamma(.)$ is

$$\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx \qquad \text{for any value } n$$

For $k > 2$, the mean and variance are given, respectively, by

$$\mu_T = 0 \quad \text{and} \quad \sigma_T^2 = \frac{k}{k-2}$$

Other Continuous Probability Distributions

– The gamma function has the following useful properties:

$$\Gamma(n) = (n-1)\Gamma(n-1)$$
$$\Gamma(n) = (n-1)! \qquad \text{for integer } n$$

– As k increases toward infinity, the variance of the t distribution approaches unity, and therefore it approaches the standard normal distribution

Other Continuous Probability Distributions

- Properties of the *t* distribution
  - It is of interest in statistical analysis to determine the percentage points $t_{\alpha,k}$ that correspond to the following probability:

$$\alpha = \mathrm{P}\left(T > t_{\alpha,k}\right) \qquad \text{or} \qquad \alpha = \int_{t_{\alpha,t}}^{\infty} f_T(t)\,dt$$

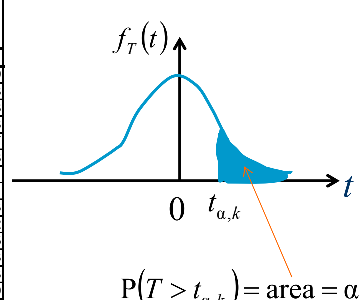
  - The percentage points are usually provided in tables. For the lower tail, the following relationship can be used:

$$t_{1-\alpha,k} = -t_{\alpha,k}$$

Other Continuous Probability Distributions

- Critical Values for *t* Distribution

Degrees of Freedom, K	Level of Significance, α				
	0.25	**0.1**	**0.05**	**0.025**	**0.005**
1	1.000001	3.077685	6.313749	12.70615	63.6559
2	0.816497	1.885619	2.919987	4.302656	9.924988
3	0.764892	1.637745	2.353363	3.182449	5.840848
4	0.740697	1.533206	2.131846	2.776451	4.60408
5	0.726687	1.475885	2.015049	2.570578	4.032117
6	0.717558	1.439755	1.943181	2.446914	3.707428
7	0.711142	1.414924	1.894578	2.364623	3.499481
8	0.706386	1.396816	1.859548	2.306006	3.355381
9	0.702722	1.383029	1.833114	2.262159	3.249843
10	0.699812	1.372184	1.812462	2.228139	3.169262
11	0.697445	1.36343	1.795884	2.200986	3.105815
12	0.695483	1.356218	1.782287	2.178813	3.054538
13	0.69383	1.350172	1.770932	2.160368	3.012283
14	0.692417	1.345031	1.761309	2.144789	2.976849
15	0.691197	1.340605	1.753051	2.131451	2.946726
16	0.690133	1.336757	1.745884	2.119905	2.920788
17	0.689195	1.333379	1.739606	2.109819	2.898232
18	0.688364	1.330391	1.734063	2.100924	2.878442
19	0.687621	1.327728	1.729131	2.093025	2.860943
20	0.686954	1.325341	1.724718	2.085962	2.845336



Other Continuous Probability Distributions

- <u>Example:</u> *Student-t* Distribution

  1. Find P($-t_{0.025,10} < T < t_{0.05,10}$)
  2. Find t_1 such that P($t_1 < T < -1.761$) = 0.045, and k = 14

Since $t_{0.05,10}$ leaves an area of 0.05 to the right and $-t_{0.025,10}$ leaves an area Of 0.025 to the left, therefore,

$$\mathrm{P}\left(-t_{0.025,10} < T < t_{0.05,10}\right) = 1 - 0.05 - 0.025 = 0.925$$

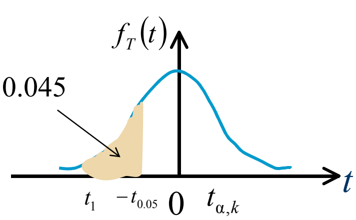
From the table, 1.761 corresponds to $t_{0.05,\,14}$ when $k = 14$
Therefore, $-t_{0.05,\,14} = -1.761$. Since t_1 in the original probability statement Is to the left of $-t_{0.05,\,14} = -1.761$, let $t_1 = -t_{\alpha,14}$. Then from the following figure we have

$$0.045 = 0.05 - \alpha$$

Other Continuous Probability Distributions

- <u>Example:</u> *Student-t* Distribution

Or α = 0.005



From the table with k =14

$$t_1 = -t_{0.005} = -2.977$$

Thus,

$$\mathrm{P}(-2.977 < T < -1.761) = 0.045$$

Other Continuous Probability Distributions

- The F Distribution

The F distribution has two shape parameters $\nu_1 = k$ and $\nu_2 = u$, and has the following PDF:

$$f_F(f) = \frac{\Gamma\left(\frac{u+k}{2}\right)\left(\frac{k}{u}\right)^{\frac{k}{2}}(f)^{\frac{k}{2}-1}}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{u}{2}\right)\left[\frac{fk}{u}+1\right]^{\frac{u+k}{2}}} \quad \text{for } f > 0$$

The mean and variance are given by

$$\mu_F = \frac{u}{u-2} \quad \text{and} \quad \sigma_F^2 = \frac{2u^2(u+k-2)}{k(u-2)^2(u-4)} \quad \text{for } u > 4$$

Other Continuous Probability Distributions

- The F Distribution
  - The distribution is positively skewed with a shape that depends on k and u.
  - It is of interest in statistical analysis to determine the percentage points $f_{\alpha,k,u}$ that correspond to the following probability:

$$\alpha = \mathrm{P}(F > f_{\alpha,k,u}) = \int_{f_{\alpha,k,u}}^{\infty} f_F(x)\,dx = \alpha$$

Other Continuous Probability Distributions

- The *F* Distribution
  - The percentage points are usually provided in mathematical book tables.
  - The *F* distribution has a unique property that allows tabulating values for the upper tail only.
  - For the lower tail, the following relation can be used to find the percentage points:

$$f_{1-\alpha,k,u} = \frac{1}{f_{\alpha,k,u}}$$

$$\text{Note: } f_{\alpha,u,k} \neq f_{\alpha,k,u}$$

Other Continuous Probability Distributions

- The *F* Distribution

Upper values for 5% (First row) and 1% (second row) Significance Level α

Second degrees of freedom, *u*	First degrees of freedom, k					
	1	2	3	4	5	6
11	4.844338	3.982308	3.587431	3.356689	3.20388	3.094613
	9.646101	7.205699	6.216737	5.668312	5.315997	5.069182
12	4.747221	3.88529	3.4903	3.25916	3.105875	2.996117
	9.330279	6.926598	5.952529	5.411948	5.064351	4.820549
13	4.667186	3.805567	3.410534	3.179117	3.025434	2.915272
	9.073801	6.70093	5.739366	5.205322	4.86159	4.620347
14	4.600111	3.73889	3.343885	3.112248	2.958245	2.847727
	8.861662	6.514938	5.563891	5.035417	4.694982	4.455842
15	4.543068	3.682317	3.287383	3.055568	2.901295	2.790465
	8.683173	6.358846	5.41695	4.893195	4.555602	4.318281
17	4.451323	3.591538	3.196774	2.964711	2.809998	2.698656
	8.399752	6.112145	5.185029	4.668948	4.335959	4.10148
20	4.35125	3.492829	3.098393	2.866081	2.710891	2.598981
	8.095981	5.84896	4.938215	4.430717	4.102674	3.871435

Other Continuous Probability Distributions

- *Chi-square* (χ^2) Distribution
  - This distribution is frequently encountered in statistical analysis, where we deal with the sum of squares of *k* random variables with standard normal distribution,

$$\chi^2 = C = Z_1^2 + Z_2^2 + \ldots + Z_k^2$$

  - Where *C* = random variable with *chi-square*, and Z_1 to Z_k are normally distributed (standard normal)

Other Continuous Probability Distributions

- *Chi-square* (χ^2) Distribution
  - The probability density function (PDF) of the *chi-square* distribution is

$$f_C(c)\frac{1}{2^{0.5k}\Gamma\left(\frac{k}{2}\right)}c^{0.5k-1}e^{\left(\frac{-c}{2}\right)} \qquad \text{for } c > 0$$

The mean and variance are given, respectively, by

$$\mu_C = k \quad \text{and} \quad \sigma_C^2 = 2k$$

Other Continuous Probability Distributions

- *Chi-square* (χ^2) Distribution
  - This distribution is positively skewed with a shape that depends on the parameter k.
  - It is of interest in statistical analysis to determine the percentage points $c_{\alpha,k}$ that correspond to the following probability:

$$\alpha = \mathrm{P}\left(C > c_{\alpha,k}\right) = \int_{c_{\alpha,k}}^{\infty} f_C(c)\,dc$$

These percentage points are usually provided in tables.

Other Continuous Probability Distributions

- Extreme Value Distributions
  - In many engineering applications, the extreme values of random variables are of special importance.
  - The largest or smallest values of random variables may dictate a particular design.
  - Wind speeds, for example, are recorded continuously at airports and weather stations. The maximum wind speeds per hour, month, day, year, or other period can be used for this purpose

Other Continuous Probability Distributions

- Extreme Value Distributions
  - Usually, the information on yearly maximum wind speed is used in engineering profession.
  - If the design wind speed has a 50-year return period, then the probability that the wind speed will exceed the design value in a year is 1/50 = 0.02.
  - Design of earthquake loads, flood levels, and so forth are also determined in this manner.

Other Continuous Probability Distributions

- Extreme Value Distributions
  - In some cases, the minimum value of a random variable is also of interest for design applications.
  - For example, when a large number of identical devices, such as calculators or cars, are manufactured, their minimum service lives are of great interest to consumers.
  - In constructing an extreme value distribution, an underlying random variable with a particular distribution is necessary.

Other Continuous Probability Distributions

- Engineering Significance of Extreme Values
  - In structural reliability and safety, the maximum loads and low structural resistance are the values most relevant to assure safety or reliability of a structure.
  - The prediction of future conditions is often required in engineering design, and may involve the prediction of the largest or smallest value.

Other Continuous Probability Distributions

- Engineering Significance of Extreme Values
  - Therefore, extrapolation from previously observed extreme value data is invariably necessary.
  - The asymptotic theory provides a powerful basis for developing the required engineering information.

Other Continuous Probability Distributions

- Type I Extreme Value Distributions
  - Two forms of the Type I extreme value distribution can be used:
    - The largest extreme value
    - The smallest extreme value

Other Continuous Probability Distributions

- The Type I Extreme Value Distribution (Largest)
  - The probability density function (PDF) of the Type I for largest distribution is

$$f_{X_n}(x) = \alpha_n e^{-\alpha_n (x-u_n)} \exp\left[-e^{-\alpha_n (x-u_n)}\right]$$

The CDF is given by

$$F_{X_n}(x) = \exp\left[-e^{-\alpha_n (x-u_n)}\right]$$

The mean and variance are given, respectively, by

$$\mu_{X_n} = u_n + \frac{\gamma}{\alpha_n} \quad \text{and} \quad \sigma^2_{X_n} = \frac{\pi^2}{6\alpha_n^2} \quad (\gamma = 0.577216)$$

Other Continuous Probability Distributions

- The Type I Extreme Value Distribution (Smallest)
  - The probability density function (PDF) of the Type I for smallest distribution is

$$f_{X_1}(x) = \alpha_1 e^{-\alpha_1 (x-u_1)} \exp\left[-e^{-\alpha_1 (x-u_1)}\right]$$

The CDF is given by

$$F_{X_1}(x) = 1 - \exp\left[-e^{-\alpha_1 (x-u_1)}\right]$$

The mean and variance are given, respectively, by

$$\mu_{X_1} = u_1 - \frac{\gamma}{\alpha_1} \quad \text{and} \quad \sigma_{X_1}^2 = \frac{\pi^2}{6\alpha_1^2} \quad (\gamma = 0.577216)$$

Other Continuous Probability Distributions

- Applications of Type I Distribution
  - Strength of brittle materials (Johnson 1953) can be described by Type I smallest value
  - Hydrological phenomena such as the maximum daily flow in a year or the annual peak flow hourly discharge during flood (Chow 1952)
  - Wind maximum velocity in a year.

Other Continuous Probability Distributions

- Example: Type I Largest

The data on maximum wind velocity V_n at a site have been compiled for n years, and its mean and standard deviation are estimated to be 61.3 mph and 7.52 mph, respectively. Assuming that V_n has a Type I extreme value distribution, what is the probability that the maximum wind velocity will exceed 100 mph in any given year?

Other Continuous Probability Distributions

- Example (cont'd): Type I Largest

The parameters u_n and α_n can be calculated as

$$\alpha_n = \sqrt{\frac{\pi^2}{6\sigma_{X_n}^2}} = \sqrt{\frac{\pi^2}{6(7.52)^2}} = 0.17055 \quad \text{and} \quad u_n = \mu_{X_n} - \frac{\gamma}{\alpha_n} = 61.3 - \frac{0.5772}{0.17055} = 57.9157$$

The probability that the maximum wind velocity is greater than 100 mph is

$$\begin{aligned} P(X_n > 100) &= 1 - F_{X_n}(x) = 1 - \exp\left[-e^{-\alpha_n(x-u_n)}\right] \\ &= 1 - \exp\left[-e^{-0.17055(100-57.9157)}\right] \\ &= 0.000763 \end{aligned}$$

Other Continuous Probability Distributions

- Example: Type I Largest

Suppose that in the previous example the design wind speed with a return period of 100 years needs to be estimated for a particular site. With V_d denoted as the design wind speed to be estimated, the probability that it will be exceeded in a given year is 1/100 = 0.01. Thus,

$$\mathrm{P}(X_n > V_d) = 1 - F_{X_n}(V_d) = 0.01$$

or

$$1 - \exp\left[-e^{-0.17055(V_d - 57.9157)}\right] = 0.01 \quad \Rightarrow \quad V_d = 84.89 \text{ mph}$$

Other Continuous Probability Distributions

- The Type II Extreme Value Distribution
  - Largest
  - Smallest
- The Type III Extreme Value Distribution
  - Largest
  - Smallest

Other Continuous Probability Distributions

- Type II
  - Largest

$$F_Y(y) = e^{-\left(\frac{v}{y}\right)^k} \quad \text{for } y \le 0$$

and

$$f_Y(y) = \frac{k}{v}\left(\frac{v}{y}\right)^{k+1} e^{-\left(\frac{v}{y}\right)^k}$$

Other Continuous Probability Distributions

- Type II
  - Largest

$$\mu_Y = v\Gamma\left(1-\frac{1}{k}\right) \quad \text{for } k>1$$

$$\sigma_Y^2 = v^2\left[\Gamma\left(1-\frac{2}{k}\right) - \Gamma^2\left(1-\frac{1}{k}\right)\right] \quad \text{for } k>2$$

$$COV = \frac{\Gamma\left(1-\frac{2}{k}\right)}{\Gamma^2\left(1-\frac{1}{k}\right)} - 1$$

Other Continuous Probability Distributions

- Gamma Function Properties

$$\Gamma(t) = \int_0^\infty r^{t-1} \exp(-r)\,dr$$

$$\Gamma(t) = (t-1)\Gamma(t-1)$$

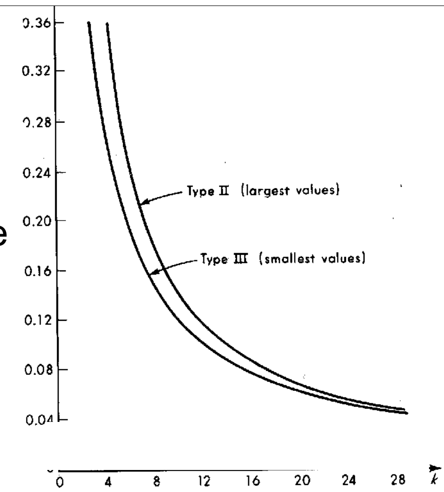
For an integer n, the gamma function becomes the factorial:

$$\Gamma(n) = (n-1)!$$

Mathematical handbooks usually contain tabulated values of the gamma function

Other Continuous Probability Distributions

- Extreme Value Distribution – Coefficient of Variation versus the parameter k
- (Benjamin and Cornell, 1970)



Extreme-value distributions. Coefficient of variation versus parameter k.

Other Continuous Probability Distributions

- Example, Type II: Wind Velocity

In Boston, Massachusetts, the measured data suggest that the mean and standard deviation of the maximum annual wind velocity are 55 mph and 12.8 mph, respectively. What is the velocity *y* which will be exceeded with a probability value of 0.02?

Other Continuous Probability Distributions

- Example, Type II (cont'd): Wind Velocity

$$COV = \frac{12.8}{55} = 0.23$$

From Fig.1, $k = 6.5$

$$\mu_Y = v\Gamma\left(1-\frac{1}{k}\right) \quad \Rightarrow \quad v = \frac{\mu_Y}{\Gamma\left(1-\frac{1}{k}\right)} = \frac{55}{\Gamma\left(1-\frac{1}{6.5}\right)} = \frac{55}{\Gamma(0.846)}$$

$$\Gamma(x+1) = x\Gamma(x) \quad \Rightarrow \Gamma(x) = \frac{\Gamma(x+1)}{x}$$

$$\Gamma(1.846) = 0.94411 \quad \Rightarrow \Gamma(0.846) = \frac{0.94411}{0.846} = 1.12$$

Other Continuous Probability Distributions

- Example, Type II (cont'd): Wind Velocity

$$\mu_Y = v\Gamma\left(1-\frac{1}{k}\right) \quad \Rightarrow \quad v = \frac{\mu_Y}{\Gamma\left(1-\frac{1}{k}\right)} = \frac{55}{\Gamma\left(1-\frac{1}{6.5}\right)} = \frac{55}{1.12} = 49.4 \text{ mph}$$

$$1 - F_Y(y) = 0.02$$

$$1 - e^{-\left(\frac{49.4}{y}\right)} = 0.02$$

$$or$$

$$y = 91 \text{ mph}$$

Other Continuous Probability Distributions

- Type II
  - Smallest

$$F_Z(z) = 1 - e^{-\left(\frac{v}{z}\right)^k} \qquad \text{for } z \le 0$$

and

$$f_Z(z) = -\frac{k}{v}\left(\frac{v}{z}\right)^{k+1} e^{-\left(\frac{v}{z}\right)^k}$$

Other Continuous Probability Distributions

- Type II
  - Smallest

$$\mu_Z = v\Gamma\left(1-\frac{1}{k}\right) \qquad \text{for } k>1$$

$$\sigma_Z^2 = v^2\left[\Gamma\left(1-\frac{2}{k}\right)-\Gamma^2\left(1-\frac{1}{k}\right)\right] \qquad \text{for } k>2$$

$$COV = \frac{\Gamma\left(1-\frac{2}{k}\right)}{\Gamma^2\left(1-\frac{1}{k}\right)}-1$$

Other Continuous Probability Distributions

- Type III
  - Largest
  - Smallest
    - Most useful applications of this model deal with smallest values.

Other Continuous Probability Distributions

- Type III
  - Smallest

$$F_Z(z) = 1 - \exp\left[-\left(\frac{z-\omega}{u-\omega}\right)^k\right] \quad \text{for } z \geq \omega$$

$$f_Z(z) = \frac{k}{u-\omega}\left(\frac{z-\omega}{u-\omega}\right)^{k-1} \exp\left[-\left(\frac{z-\omega}{u-\omega}\right)^k\right] \quad \text{for } z \geq \omega$$

Other Continuous Probability Distributions

- Type III
  - Smallest

$$\mu_Z = \omega + (u-\omega)\left[\Gamma\left(1+\frac{1}{k}\right)\right]$$

$$\sigma_Z^2 = (u-\omega)^2\left[\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)\right]$$