ISyE 6739 –Statistical Methods Review of Probability (Chapters 3, 4 & 7)

Instructor: Kamran Paynabar
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

Kamran.paynabar@isye.gatech.edu
Office: Groseclose 436

ISyE 6739, Probability Review

1

List of Topics

- Review of Probability
- Discrete Distributions
 - Binomial
 - Poisson
 - Geometric
 - Negative-Binomial
- Continuous Distributions
 - Exponential
 - Normal
 - Chi-square
 - Student's t
 - F

Review of Probability Distributions

ISyE 6739, Probability Review

3

Random Variables

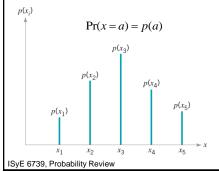
- Random Experiment: An experiment that can result in different outcomes, even
 if repeated in the same manner every time, e.g., rolling a dice, producing a part.
- Random Variable: A variable that associates a number with the outcome of a random experiment is called a random variable.
 - X: Compressive strength of each specimen
 - Y: Number of defectives in a sample of *n* produced parts
- Type of RVs:
 - A discrete random variable is a rv with a finite (or countably infinite) range. They
 are usually integer counts, e.g., number of errors or number of bit errors per
 100,000 transmitted (rate).
 - A continuous random variable is a rv with an interval (either finite or infinite) of real numbers for its range. Its precision depends on the measuring instrument.

Discrete Distributions

Discrete distributions. When the parameter being measured can only take on certain values, such as the integers $0, 1, 2, \ldots$, the probability distribution is called a *discrete distribution*.

Examples: Quality of a part (Ok = 0, Not ok = 1), Number of typos on a typed page (0,1,2,...)

Probability <u>mass</u> function (pmf): p(x) is pmf \leftrightarrow $\begin{cases} p(x) \ge 0 \\ \sum_{-\infty}^{+\infty} p(x) = 1 \end{cases}$



Examples: X:Number of observed 6 in 2 rolls of a dice

а	0	1	2		
Pr(X=a)	(5/6)2	2(1/6)1(5/6)1	(1/6)2		

$$\Pr(X = a) = {2 \choose a} \left(\frac{1}{6}\right)^a \left(\frac{5}{6}\right)^{2-a}; \ a = 0,1,2$$

5

Cumulative Distribution Function (CDF)

Cumulative Distribution Function (CDF): The CDF of a discrete random variable X is denoted by F(X=a) and is calculated by

$$F(X = a) = F(a) = \Pr(x \le a) = \sum_{i = -\infty}^{a} p(i)$$

Some Properties:

- 1) $0 \le F(a) \le 1$
- 2) if $x \le y$, then $F(x) \le F(y)$
- 3) $Pr(a < x \le b) = Pr(x \le b) Pr(x \le a) = F(b) F(a)$
- 4) $\Pr(a \le x \le b) = \Pr(x \le b) \Pr(x \le a 1) = F(b) F(a 1)$

Continuous Distributions

Continuous distributions. When the variable being measured is expressed on a continuous scale, its probability distribution is called a *continuous distribution*.

Examples: Diameter of piston rings, viscosity of oil, customers' Waiting time

Probability density function (pdf):

f(x) is pdf $\leftrightarrow \begin{cases} f(x) \ge 0 \\ \int_{-\infty}^{+\infty} f(x) dx = 1 \end{cases}$

 $\Pr(a \le x \le b) = \int_{a}^{b} f(x)dx$ $\Pr(x = a) = 0 \neq f(a)$

 $F(X = a) = F(a) = \Pr(x \le a) = \int_{-\infty}^{a} f(x)dx$

7

Cumulative disruption function (cdf):

ISyE 6739, Probability Review

Example: The probability distribution of X is

$$f(x) = 2e^{-kx}, x \ge 0;$$

- (a) Find the appropriate value of k;
- (b) Find the Pr(X>0.5).

Example: The product quality characteristic measurement is a random variable X with probability distribution

$$f(x) = \begin{cases} 4(x-11.75) & 11.75 \le x \le 12.25 \\ 4(12.75-x) & 12.25 < x \le 12.75 \end{cases}$$

If the quality specification limit is that X must be less than12.50, what is the percentage of products to be nonconforming?

- 0.125 a)
- b) 0.5
- c) 0.875

ISVE 6739, Propability Review

Mean, Variance and Moments

• Mean is a measure of the center of a probability distribution.

$$E(x) = \mu = \sum_{n=0}^{+\infty} xp(x)$$
 Discrete RV

$$E(x) = \mu = \sum_{-\infty}^{+\infty} x p(x)$$
 Discrete RVs
$$E(x) = \mu = \int_{-\infty}^{+\infty} x f(x) dx$$
 Continuous RVs

Moments are measures related to the location/shape of probability distribution.

$$E(x^r) = \sum_{-\infty}^{+\infty} x^r p(x)$$
 Discrete RVs

$$E(x^r) = \int_{-\infty}^{+\infty} x^r f(x) dx$$
 Continuous RVs

Variance is a measure of the dispersion or variability of a probability distribution.

$$V(x) = \sigma^2 = \sum_{n=0}^{+\infty} (x - \mu)^2 p(x)$$
 Discrete RVs

$$V(x) = \sigma^2 = \sum_{-\infty}^{+\infty} (x - \mu)^2 \, p(x) \qquad \text{Discrete RVs}$$

$$V(x) = \sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 \, f(x) dx \qquad \text{Continuous RVs}$$

$$V(x) = E(x^2) - [E(x)]^2$$

ISyE 6739, Probability Review

Example: Find mean and variance for the following RVs

a) X: Number of observed 6 in 2 rolls of a dice

а	0	1	2		
Pr(X=a)	(5/6)2	2(1/6)1(5/6)1	(1/6)2		

b) X: exam time limited between 0 and 1 hour.

$$f(x) = \alpha(x)^{\alpha-1}; 0 < x < 1$$

ISyE 6739, Probability Review

11

Useful Results on Mean and Variance

If x is a random variable and a is a constant, then

$$E(a+x)=a+E(x)$$

$$E(a \times x) = aE(x)$$

$$V(a+x)=V(x)$$

$$V(a \times x) = a^2 V(x)$$

If $x_1, x_2, ..., x_n$ are random variables,

$$E(a_1x_1+...+a_nx_n)=a_1E(x_1)+...+a_nE(x_n)$$

If they are **mutually independent**, and $a_1,...,a_n$ are constants

$$V(a_1X_1 + ... + a_nX_n) = a_1^2V(X_1) + ... + a_n^2V(X_n)$$

12

ISyE 6739, Probability Review

Review of Discrete Distributions

ISyE 6739, Probability Review

13

Bernoulli Distribution

 $X \sim Bernoulli(p)$ x = 0,1 p: probability of success

Application: To model the outcome of a Bernoulli trial with two possible outcomes (Success = 1, Failure = 0).

Example: The outcome of inspecting part quality;

Nonconforming (Success=1), Conforming (Failure=0)

pmf, mean and variance:

$$P(x) = p^{x}(1-p)^{1-x}; x = 0,1$$

$$E(x) = p$$

$$Var(x) = p(1-p)$$

ISyE 6739, Probability Review

Binomial Distribution

 $X \sim Bin(n, p)$ x = 0,1,...,n p: probability of success

Application: To model the number of successes in *n* independent Bernoulli trials.

Example: number of nonconforming parts in a random sample of parts of size *n*.

pmf, mean and variance:

$$P(x) = \binom{n}{x} p^{x} (1-p)^{n-x}; \ x = 0,1,...,n$$

$$E(x) = np$$
 $Var(x) = np(1-p)$

Fact: sum of n independent Bernoulli variables (p) is Binomial (n,p)

ISyE 6739, Probability Review

15

Examples

- (1) In a welding process, 10 different points of a battery tab are welded. If the probability of making an acceptable weld is 0.99, calculate:
 - a) the probability of making a battery tab with 6 failed welds.
 - b) the probability of making a battery with at least 2 failed welds.

Poisson Distribution

$$X \sim Poisson(\lambda)$$
 $x = 0,1,2,...$ λ : average number of events in a unit (a.k.a. rate)

Application: To model the number of events in a desired unit

Example: number of scratches on the surface of a refrigerator.

events unit

pmf, mean and variance:

$$P(x) = \frac{\lambda^x \exp(-\lambda)}{x!}; \quad x = 0,1,2,\dots \qquad E(x) = Var(x) = \lambda$$

Fact: sum of *n* independent Poisson variables (λ) is Poisson with the rate of ($n\lambda$)

ISyE 6739, Probability Review

17

Examples

- (2) In a painting shop, on average, there are 5 defects on the body of a painted car. Calculate
 - a) the probability of finding a painted body with 10 defects.
 - b) the probability of finding 10 defects on 2 painted bodies.

Geometric Distribution

$$X \sim Geo(p)$$
 $x = 1, 2, ...$

p: probability of success

Application: To model the number of Bernoulli trails until the first success

Example: number of inspected part until the first nonconforming part is found

pmf, mean and variance:

$$P(x) = p(1-p)^{x-1}; x = 1,2,...$$

$$E(x) = \frac{1}{p}$$

$$Var(x) = \frac{1}{p} \left(\frac{1}{p} - 1 \right)$$

Fact: the number of Bernoulli trails between two successes also follows a Geometric distribution (because of Memoryless property)

ISyE 6739, Probability Review

Negative-Binomial (Pascal) Distribution

$$X \sim NB(r, p)$$
 $x = r, r + 1, \dots$ p : probability of success

Application: To model the number of Bernoulli trails until *r* successes

Example: number of inspected part until the *r*th nonconforming part is found

pmf, mean and variance:

pmf, mean and variance:
$$E(x) = \frac{r}{p}$$

$$P(x) = {x-1 \choose r-1} p^r (1-p)^{x-r}; \ x = r, r+1, \dots$$

$$Var(x) = \frac{r}{p} \left(\frac{1}{p} - 1\right)$$

Fact: sum of *r* independent Geometric variables (*p*) is Negative-Binomial

ISyE 6739, Probability Review

Examples

In a welding process, probability of making an acceptable weld is 0.99, calculate:

- a) the probability of making at least 5 welds until a defective weld is made
- b) the probability of making the 3rd defective weld in the 5th battery tab.

ISyE 6739, Probability Review

21

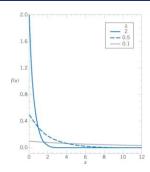
Review of Continuous Distributions

Exponential Distribution

$$X \sim Exp(\lambda)$$
 $x \ge 0$

$$f(x) = \lambda \exp(-\lambda x); x \ge 0$$

Example: waiting time of customers, time to failure of a machine



CDF, mean and variance:

$$E(x) = \frac{1}{\lambda}$$

$$Var(x) = \frac{1}{\lambda^2}$$

$$F(a) = \Pr(X \le a) = 1 - \exp(-a\lambda)$$

Memoryless property: Pr(x>a+b|x>a) = Pr(x>b)

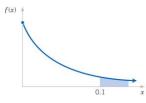
ISyE 6739, Probability Review

23

Example 4-21: Computer Usage-1

In a large corporate computer network, time to user log-ons to the system can be modeled as exponential distribution with a mean of 0.04 hours. What is the probability that there are no log-ons in the next 6 minutes (0.1 hours)?

Let *X* denote the time in hours from the start of the interval until the first log-on.



what is the interval of time such that the probability that no log-on occurs during the interval is 0.90?

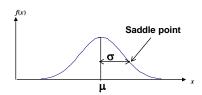
Normal Distribution

$$X \sim N(\mu, \sigma^2); \quad -\infty < x < +\infty$$

$$1 \quad -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)$$

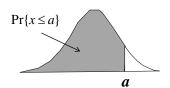
$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}$$

$$E(x) = \mu$$
 $Var(x) = \sigma^2$



Fact: If x_1 , x_2 are independently normally distributed variables, then y=x₁+x₂ also follows the normal distribution, i.e. y~N($\mu_1+\mu_2$, $\sigma_1^2+\sigma_2^2$)

$$\Pr\{x \le a\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^{2}\right\} dx$$



ISyE 6739, Probability Review

25

Standard Normal Distribution

$$X \sim N(\mu, \sigma^2); \quad -\infty < x < +\infty$$

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

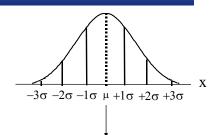
Map any X into Z

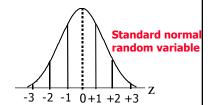
$$Z = \frac{X - \mu}{\sigma}$$

$$Z \sim N(0,1^2); \quad -\infty < z < +\infty$$

$$f(z; \mu = 0, \sigma^2 = 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

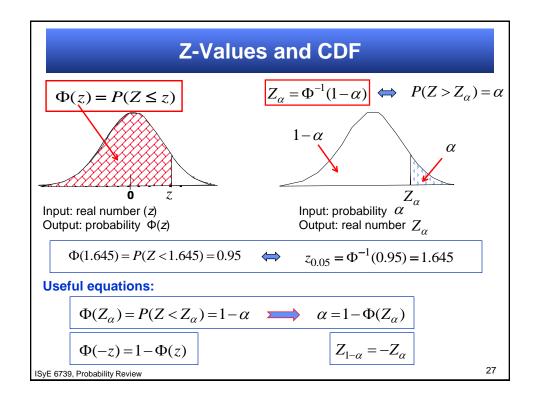
Excel Function: $F(x)=NORMDIST(x,\mu,\sigma,true)$

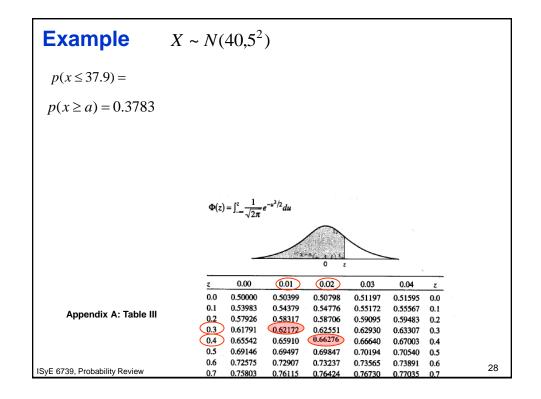




 $Pr(\mu - \sigma \le x \le \mu + \sigma) = 68.26\%$ $Pr(\mu-2\sigma \leq x \leq \mu+2\sigma)=95.46\%$ $Pr(\mu - 3\sigma \le x \le \mu + 3\sigma) = 99.73\%$

ISyE 6739, Probability Review





Example: Three shafts are made and assembled in a linkage. The length of each shaft, in centimeters, is distributed as follows:

> Shaft 1: N ~ (75, 0.09) Shaft 2: N ~ (60, 0.16) Shaft 3: N ~ (25, 0.25)

Assume the shafts' length are independent to each other:

- (a) What is the distribution of the linkage?
- (b) What is the probability that the linkage will be longer than 160.5 cm?

ISyE 6739, Probability Review

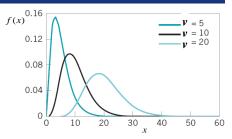
29

Chi-square Distribution

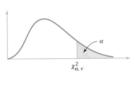
$$X \sim \chi^{2}(\nu) \qquad E(X) = \nu$$
Degrees of freedom
$$Var(X) = 2\nu$$

$$f(x;\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \exp\left(-\frac{x}{2}\right) x^{\left(\frac{\nu}{2}-1\right)}$$

E(X) = v



The Chi-square pdf is not easily integrable (except for v=2). The CDF table (Appendix A, Table IV) is usually used.



 $\Gamma(a) = (a-1)!; a \in Z^+$

, a	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	+00.	+00.	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19

ISyE 6739, Probability Review

Chi-square Distribution

• Fact 1: If $Z_1, ..., Z_k$ are *independent*, *standard normal* random variables, then the sum of their squares is distributed according to the chi-square distribution with k degrees of freedom.

$$X = \sum_{i=1}^k Z_i^2 \sim \chi^2(k)$$

• Fact 2: If X_i's (i = 1,2,...,n) are *independent*, *Chi-square* random variables with v_i's degrees of freedom, then

$$Y = \sum_{i=1}^{n} X_i \sim \chi^2(\sum_{i=1}^{n} v_i)$$

ISyE 6739, Probability Review

31

Example

- Calculate the following quantities:
 - 1. $\chi^2_{0.05.5}$

$$2. \Pr \left\{ \sum_{i=1}^{8} \left(\frac{X_i - \mu}{\sigma} \right)^2 > 1.34 \right\}$$

Student's t Distribution

$$f(t;v) = \frac{1}{\sqrt{\pi v}} \frac{\Gamma(\frac{v+1}{2})}{\Gamma(v/2)} \left(1 + \frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)} = 0.4$$

$$E(t) = 0$$

$$Var(t) = \frac{v}{v-2}; v > 2$$

$$0.4$$

$$0.4$$

$$0.3$$

$$0.3$$

$$0.4$$

$$0.3$$

$$0.4$$

$$0.3$$

$$0.3$$

$$0.4$$

$$0.4$$

$$0.3$$

$$0.4$$

$$0.4$$

$$0.3$$

$$0.4$$

$$0.4$$

$$0.4$$

$$0.4$$

$$0.7$$

$$0.4$$

$$0.7$$

$$0.4$$

$$0.7$$

$$0.4$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.7$$

$$0.$$

The student's t pdf is not easily integrable (except for v = 1). The CDF table (Appendix A, Table V) is usually used.

Table V Percentage Points $t_{\alpha,\nu}$ of the t Distribution .0005 63.657 127.32 318.31 636.62 12.706 31.821 .325 1.000 3.078 6.314 31.598 6.065 9.925 14 089 23.326 12.924 7.453 .277 4.541 5.841 .765 1.638 2.353 3.182 2.132 .741 1.533 2.571 3.365 4.032 4.773 5.893 6.869 1.440 1.943 2.447 3.143 3.707 4.029 4.785 5.408 2.365 .263 .711 1.415 5.041 2.306 2.896 3.355 3.833 3.250 .261 .703 1.383 1.833 2.262 2.821 ISyE 6739, Probability Review

Student's t Distribution

- Fact 1: If
 - Z is a standard normal (N(0,1)) random variable
 - V has a chi-square distribution with v degrees of freedom;
 - Z and V are independent.

Then the following ratio follows a student's t

$$\frac{Z}{\sqrt{V/\nu}} \sim t(\nu)$$

• Fact 2: the student's t distribution tends to standard normal distribution as *v* increases (*v* > 30)

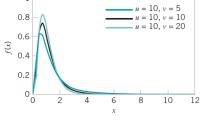
Example

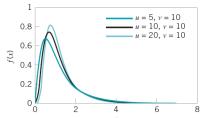
- Calculate the following quantities:
- $1.t_{0.05,5}$
- 2. $\Pr(\frac{Z}{\sqrt{\left(\frac{X-\mu_x}{\sigma_x}\right)^2}} > 1)$ Z follows a standard normal dist. and X follows a normal dist. and they are independent

ISyE 6739, Probability Review

35







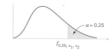
$$f(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)} \frac{x^{(u/2)-1}}{\left[\left(\frac{u}{v}\right)x+1\right]^{(u+v)/2}}, 0 < x < \infty$$

ISyE 6739, Probability Review

F Distribution

The *F* pdf is not easily integrable. The CDF table (Appendix A, Table VI) is usually used.

Table VI Percentage Points $f_{\alpha,\nu_{\alpha},\nu_{\gamma}}$ of the F Distribution



									30.23	11112			
\ V1								Degrees of freedom for the numer				ator (v ₁)	
1/2	1	2	3	4	5	6	7	- 8	9	10	12	15	20
1	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.41	9.49	9.58
2	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.41	3,43
3	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.46	2.46
4	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
5	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88
6	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.76	1.76
7	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.70	1.69	1.68	1.68	1.67
8	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61
9	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.57	1.56
10	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.53	1.52

- A random variate of the F-distribution arises as the ratio of two chi-squared variates: $\frac{U_1/d_1}{U_2/d_2}$
 - U_1 and U_2 have *chi-square distributions* with d_1 and d_2 degrees of freedom respectively, and
 - U_1 and U_2 are independent.

ISyE 6739, Probability Review

37

Example

- Calculate the following quantities:
- $1.\,F_{0.25,6,8}$

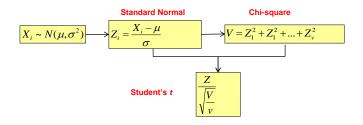
$$2.\Pr\left(\frac{\sum_{i=1}^{2} \left(\frac{X_{i} - \mu_{x}}{\sigma_{x}}\right)^{2}}{\sum_{i=1}^{2} \left(\frac{Y_{i} - \mu_{y}}{\sigma_{y}}\right)^{2}} > 3\right)$$

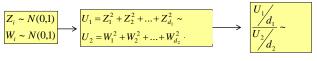
X and Y are independent and follow a normal dist.

Excel Function: f(f0)=fdist(f0,df1,df2)

ISyE 6739, Probability Review

Summary





F-distribution

ISyE 6739, Probability Review

39

Useful Fact About Sampling Distributions

Fact 1: A Chi-square random variable with 2 dof is equivalent to an exponential random variable with $\lambda = 0.5$.

$$f(x) = \lambda e^{-\lambda x}, x \ge 0;$$

Fact 2: A t random variable with 1 dof is equivalent to a Cauchy random variable

 $f(t) = \frac{1}{\pi} \times \frac{1}{1+t^2}$

Fact 3: An
$$F$$
 random variable with 2 and 2 dof's
$$f \sim F_{2,2} \rightarrow g(f) = \frac{1}{\left(1+f\right)^2}$$

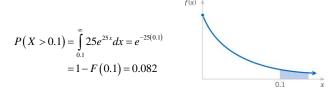
- Fact 4: if $f \sim F(v_1, v_2)$, then $1/f \sim F(v_2, v_1)$ $F_{1-\alpha, v_1, v_2} = \frac{1}{F_{\alpha, v_1, v_2}}$
- $F_{\alpha,1,\nu} = t_{\alpha/2,\nu}^2$ Fact 5: if $t \sim \text{students'} - t(v)$, then $t^2 \sim F(1, v)$

ISyE 6739, Probability Review

Example 4-21: Computer Usage-1

In a large corporate computer network, time to user log-ons to the system can be modeled as exponential distribution with a mean of 0.04 hours. What is the probability that there are no log-ons in the next 6 minutes (0.1 hours)?

Let X denote the time in hours from the start of the interval until the first log-on.



what is the interval of time such that the probability that no log-on occurs during the interval is 0.90?

$$P(X > x) = e^{-25x} = 0.90, -25x = \ln(0.90)$$

 $x = \frac{-0.10536}{-25} = 0.00421 \text{ hour} = 0.253 \text{ minute}$

ISyE 6739, Probability Review