ECE3075 - Random Signals

Chapter 2: Random Variables

Chin-Hui Lee

School of Electrical and Computer Engineering

Georgia Institute of Technology

Atlanta, GA 30332, USA

chl@ece.gatech.edu

Course Information

- **Subject:** Random Signal and Systems
- **Prerequisite:** ECE2025
- **Background Expected**
  - Fundamentals on Calculus, Physics, and Digital Signal Processing
- **Grades: HW (25%), Two Quizzes (15+20%), Final (40%)**
- **Tools Expected**
  - Basic analytical skills
  - Textbooks and reading assignments: your main source of learning
  - Class Lectures: exploring beyond the textbooks
  - Homework: hand-on and get-your-hands-dirty exercises
- **Website:** http://users.ece.gatech.edu/~chl/ECE3075.sum04/
  - Check frequently for class notes, updates, and homework assignments

Review of Chapter 1

- Deterministic vs. Probabilistic Approach
- Probability: a measure of how likely an event is
- Relative Frequency vs. Axiomatic Approach
  - –Relative Frequency Approach: attach a physical meaning
  - –Axiomatic Approach: 3 fundamental axioms to be satisfied
    1. $P(A) \geq 0 \quad \forall A \in \Omega \ (A \subset S)$
    2. $P(S) = 1 \quad S \in \Omega$ (certain event)
    3. If $A \cap B = \phi$ then $P(A \cup B) = P(A) + P(B), \forall A, B \in \Omega$

Review of Chapter 1

- Joint Probability: $P(A,B)$
- Conditional Probability: $P(A \mid B) = P(A,B) / P(B)$
- Total Probability: $P(B) = P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)$
- A Priori Probability: $P(A_i)$
- A Posterior Probability: $P(A_i|B)$
- Bayes' Rule: $P(A_i \mid B) = P(B \mid A_i)P(A_i) / P(B)$

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{[P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)]}$$

- Statistical Independence: $P(A \cap B) = P(A) * P(B)$

Review of Chapter 1

- Elementary Set Theory
- Bernoulli Trials: Binary Events

$$P(A) = P(\text{"}success\text{"}) = p,\ P(\bar{A}) = P(\text{"}failure\text{"}) = q = 1 - p$$

$$p_n(k) = P(k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k q^{(n-k)}$$

- DeMoivre-Laplace Theorem:

  If $npq \gg 1$ and $|k - np|$ is in the order of $\sqrt{npq}$ then

$$p_n(k) = \binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} \exp[-\frac{(k-np)^2}{2npq}]$$

(due to Gaussian Approximation)

Random Variables

- Just a mapping of events onto real line
- Random in the sense that the outcome cannot be predicted with certainty before the experiment

EX 1: We flip a (fair) coin

$$P(\text{"}Head\text{"}) = p,\; P(\text{"}Tail\text{"}) = q = 1 - p$$

X = Outcome of coin flipping experiment

$$X = \begin{cases} 1, \text{"}Head\text{"} \\ 0, \text{"}Tail\text{"} \end{cases} \Rightarrow X = \begin{cases} 1, \text{ with probability p} \\ 0, \text{ with probability q} \end{cases}$$

Random Variables and Vectors

- Random vector $\underline{X} = (X_1, \ldots, X_n)$ is a sequence of n random variables

EX 2: We flip a pair of coins

X = Outcome of first coin, Y = Outcome of second coin

$\underline{Z} = (X,Y)$, $V = X \oplus Y$ (Are they different or same?)

EX 3: We roll a pair of dice

$X =$ Score on the 1^{st} die, $Y =$ Score on 2^{nd} die

$\underline{Z} = (X,Y)$, $U = X + Y$, $V = \max(X, Y)$

Random Variables and Vectors

- Random Variable: A function X(.) that maps the sample space *S* to a *one*-dimensional space of real numbers for easy mathematical manipulation (sample space can be irregular)
  - linking events to numerical values $X : S \to \Re$
- Random Vector: A function X(.) that maps the sample space *S* to an *n*-dimensional space of real numbers for easy manipulation (sample space can be irregular) $X : S \to \Re^n$

Two Types of Random Variables

- Continuous (Value) Random Variable
  - link uncountable events to continuous numerical values
  - EX: resistance values, lifetime of a bulb
  - e.g.Uniform, Gaussian r.v.'s

$$\int_{-\infty}^{\infty} p(X=x)dx = P(S) = 1,\ P(X=x) = 0$$

- Discrete (Value) Random Variable
  - link countable events to discrete numerical values
  - EX: flip of a coin, roll of a die
  - e.g. Binomial and multinomial r.v.'s

$$\sum_i p(X=x_i) = \sum_i P(A_{x_i}) = P(S) = 1$$

- Same mathematical treatment

Distribution Functions

- Probability Distribution Function (PDF) of a r.v. X

$$F_X(x) = P(X \le x)$$

where X is a random variable, x is any allowable value

- Some Important Properties
  1. $0 \le F_X(x) \le 1, \ -\infty < x < \infty$
  2. $F_X(-\infty) = 0, \text{ and } F_X(\infty) = 1$
  3. $F_X(x)$ is a non-decreasing function of x
  4. $P(x_1 \le X < x_2) = F_X(x_2) - F_X(x_1)$
- Illustration Examples
  - Continuous random variables: Figure 2-2 (a) and (b)
  - Discrete random variable: Figure 2-2(c) discontinuities

Center of Signal and Image Processing
Georgia Institute of Technology
CSIP

Classroom Illustrations (I): PDF

- Blackboard Illustrations
  1. Section 2-2
  2. Exercise 2-2.1
  3. Exercise 2-2.2

Density Functions

- Probability Density Function (pdf) of a r.v. *X*, if exists

$$f_X(x) = \lim_{e\to 0}[F_X(x+e) - F_X(x)]/e = dF_X(x)/dx$$

where X is a random variable, x is any allowable value

- Some Important Properties
  1. $f_X(x) \geq 0, \quad -\infty < x < \infty$
  2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$
  3. $F_X(x) = \int_{-\infty}^{x} f_X(x)dx$
  4. $P(x_1 \leq X < x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)dx$
- Illustration Examples: Figure 2-5

Density Functions

- Discrete pdf: singularity with probability mass (delta)

$\Rightarrow$ We can write $f_X(x) = \sum_{i=1}^{n} f_X(x_i)\delta(x - x_i)$ with $f_X(x_i) = P(X = x_i)$

- Hence, for discrete random variables:

1. $f_X(x) \geq 0, \quad -\infty < x < \infty$
2. $\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} \left[\sum_{i=1}^{n} f_X(x_i)\delta(x - x_i) \right] dx = \sum_{i=1}^{n} f_X(x_i) = 1$
3. $F_X(x)\Big|_{x=x_k} = \int_{-\infty}^{x_k} f_X(x)dx = \sum_{i=1}^{k} f_X(x_i)$
4. $P(x_1 \leq X < x_2) = F_X(x_2) - F_X(x_1)$

- REMARK: For a continuous random variable $P(X = x) = 0$. WHY?

Center of Signal and Image Processing
Georgia Institute of Technology

Random Functions

- Function of a single random variable: $Y = g(X)$

$$f_X(x)dx = f_Y(y)dy$$

- pdf of random Functions

$$f_Y(y) = f_X(x) * \frac{dx}{dy} = \frac{1}{|A|} f_X(g^{-1}(y)) \text{ with } \frac{dy}{dx} = A$$

- Illustration Example: Figure 2-9 $Y = g(X) = X^2$

$$f_Y(y) = [f_X(\sqrt{y}) + f_X(-\sqrt{y})] / 2\sqrt{y},\ y \geq 0$$

- More Illustration Examples
  - Exercises 2-3.1, 2-3.2 and many others

Expectation of Random Functions

- Mathematical Expectation: $E_X[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$
- Mean of random variables: center of mass

$$\bar{X} = \mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- Moment Function $E[(X - \bar{X})^n]$
- Variance: measure of distribution spread

$$\sigma_X^2 = VAR(X) = E[(X - \bar{X})^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$$

$$= E[X^2] - 2E[X]\bar{X} + \bar{X}^2 = E[X^2] - \bar{X}^2$$

- Illustration Examples:
  - Figure 2-10, Exercise 2-4.1, 2-4.2, and many others

Center of Signal and Image Processing
Georgia Institute of Technology
CSIP

Expectation of Random Functions

- For discrete random variables: $f_X(x) = \sum_{i=1}^{n} f_X(x_i)\delta(x - x_i)$

  Hence, mean and variance are:

$$\overline{X} = \mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \sum_{i=1}^{n} x_i f_X(x_i)$$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx = \sum_{i=1}^{n} (x_i - \mu_X)^2 f_X(x_i)$$

- Some useful properties of expectation:
  1. $E[aX + b] = aE[X] + b$, a and b are real numbers
  2. $E[X + Y] = E[X] + E[Y]$
  3. If X and Y are independent, $E[X * Y] = E[X] * E[Y]$
- Exercise: Show properties 1-3

Random Variable Recap

- Continuous and Discrete Random Variable
  - mapping events to a subset of real numbers
- Probability Density Function (pdf)

$$P(a \le x \le b) = \int_a^b f(x)dx, \quad P(S) = \int_{-\infty}^{\infty} f(x)dx = 1$$

- Probability Distribution Function (PDF)

$$F(y) = \int_{-\infty}^{y} f(x)dx \quad F(-\infty) = 0 \quad F(\infty) = 1$$

- Expectation of Random Functions

$$E(q(X)) = \int_{-\infty}^{\infty} q(x)f(x)dx \quad \text{or} \quad \sum_i q(x_i)p(x_i)$$

- Mean and Variance

$$\text{Mean}(X) = \text{E}(X) \quad \text{Var}(X) = \text{E}([X - \text{E}(X)]^2)$$

Some Useful Distributions (I)

- Binomial Distribution: *B(R=r; n, p)*
  - probability of *r* successes in *n* trials with a success rate *p*

$$B(r;n,p)=\frac{n!}{r!(n-r)!}p^r(1-p)^{n-r} \quad \text{where} \quad 0\le r\le n$$

- Multinomial Distribution

$$M(r_1,\ldots,r_m;n,p_1,\ldots,p_m)=\frac{n!}{r_1!\cdots r_m!}\prod_{i=1}^{m}p_i^{r_i} \quad 0\le r_i \quad \sum_{i=1}^{m}r_i=n$$

- Show

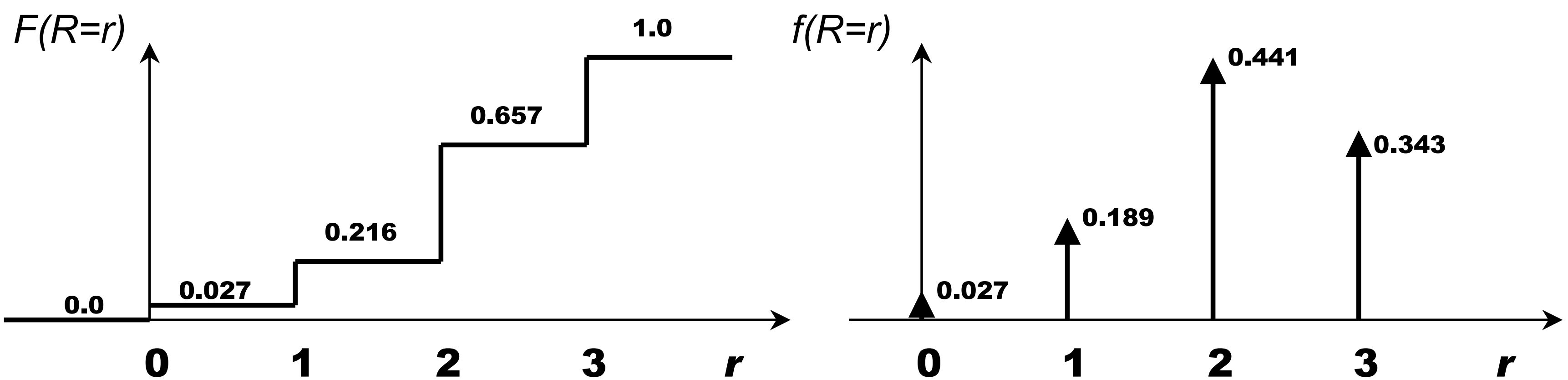
$$\sum_{r=0}^{n}B(r;n,p)=1, \quad \mathrm{E}_B(R)=\sum_{r=0}^{n}rB(r;n,p)=np, \quad VAR_B(R)=npq$$

Center of Signal and Image Processing
Georgia Institute of Technology
CSIP

PDF and pdf of Binomial Distribution

- Binomial distribution: n=3, p=0.7

$$B(r;n,p)=\frac{n!}{r!(n-r)!}p^r(1-p)^{n-r} \quad \text{where} \quad 0\le r\le n$$



- Can you plot for other value pairs of (n, p) ? You can look up the Binomial Coefficients in App. C.

Some Useful Distributions (II)

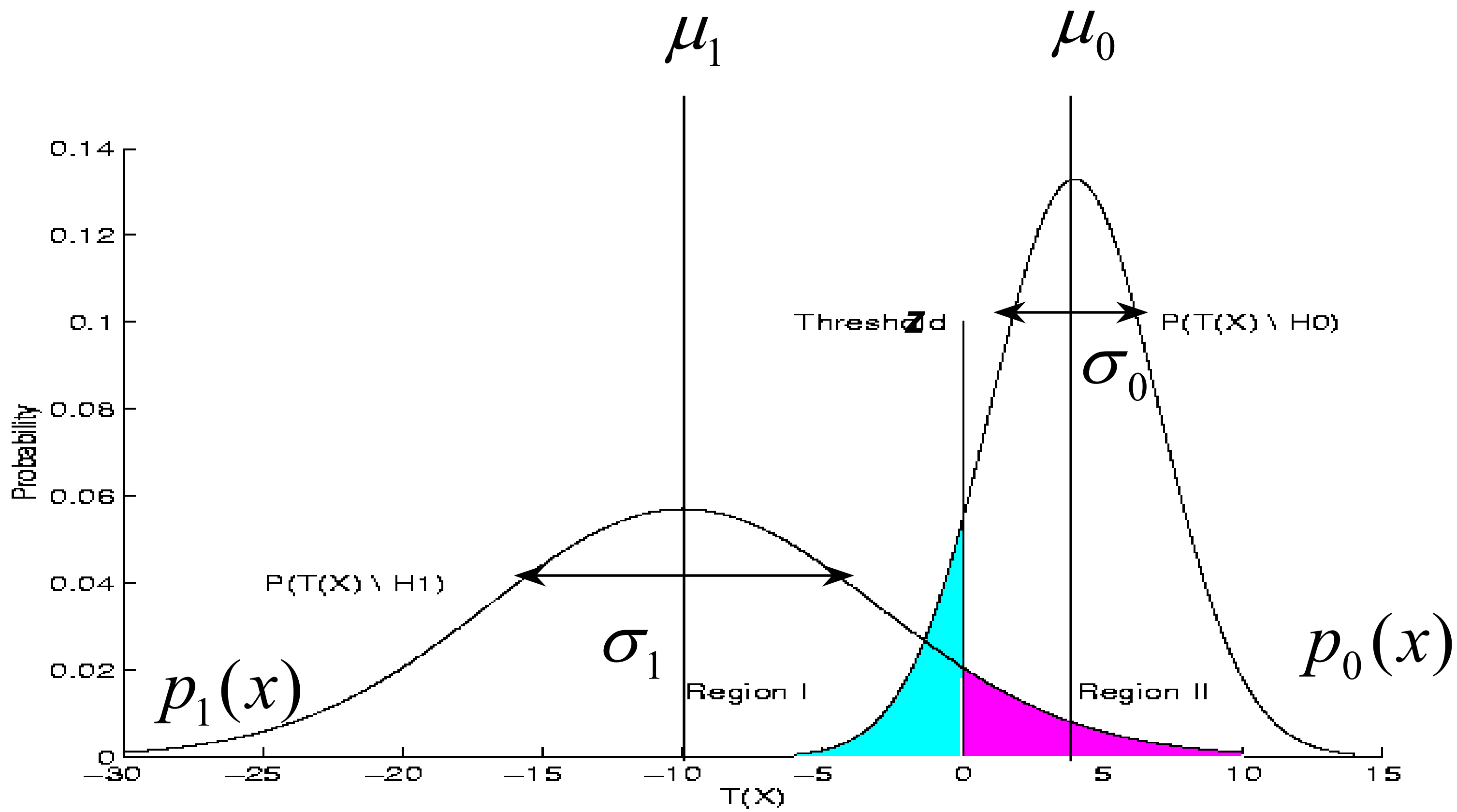
- *Normal* (or *Gaussian*) Distribution: *Bell Curve*
  - The most widely used and most studied density

$$f_N(x;\mu,\sigma^2) = N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad -\infty < x < \infty \quad \sigma > 0$$

- Show its mean: $\mathrm{E}_N(X) = \mu$
- Can you compute its variance? $Var_N(X)$
- Illustration: Section 2-5, Figure 2-11

Center of Signal and Image Processing
Georgia Institute of Technology
CSIP

Typical Gaussian Distributions



Standard deviation (s.d. or spread): $\sigma_0 < \sigma_1$

Properties of Gaussian R.V.'s

1. $\max_x f_N(x;\mu,\sigma^2) = f_N(\mu)$, single maximum, mode=mean
2. The pdf is symmetric around the mean
3. Show that:
$$f_N(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}, \text{ and } f_N(\mu \pm \sigma) = \frac{0.607}{\sqrt{2\pi\sigma^2}}$$
4. Standard deviation is a good spread measure
$$P(\mu - m\sigma < x < \mu + m\sigma) = \int_{-m\sigma}^{m\sigma} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{x^2}{2\sigma^2}]dx$$
5. Evaluate the value for m=1, 2, 3 (App. D)

Properties of Gaussian R.V.'s (Cont.)

6. PDF: $$F_N(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} \exp[-\tfrac{1}{2\sigma^2}(y-\mu)^2]dy$$

7. Equivalent to zero mean, unity variance r.v. (App. D)

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp[-\tfrac{1}{2}y^2]dy, \; F(x) = \Phi(\frac{x-\mu}{\sigma})$$

8. Other variation (App. E and G): $Q(x) = 1 - \Phi(x)$

9. Error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x} e^{-u^2}du, \quad Q(x) = \frac{1}{2}[1 - \operatorname{erf}(\frac{x}{\sqrt{2}})]$$

Center of Signal and Image Processing
Georgia Institute of Technology

Properties of Gaussian R.V.'s (Cont.)

10. Bounds on Q-function (MATLAB, App. G):

$$(1-\frac{1}{a^2})\frac{1}{a\sqrt{2\pi}}e^{-a^2/2} \le Q(a) \le \frac{1}{a\sqrt{2\pi}}e^{-a^2/2}$$

11. Approximation of Q-function (Figures 2-12, 2-13)

$$Q(a) \cong (1-\frac{1}{2a^2})\frac{1}{a\sqrt{2\pi}}e^{-a^2/2}$$

12. Higher Moment: Show

$$E[(x-\mu)^n] = \begin{cases} 0 & n \text{ odd} \\ 1*3*\cdots*(n-1)\sigma^n & n \text{ even} \end{cases}$$

13. More Illustration:
    – Section 2-5

Center of Signal and Image Processing
Georgia Institute of Technology

CSIP

Some Notes on MATLAB

- MATLAB is becoming a popular programming tool for computing and using mathematical functions
- Visualize data with preprogrammed functions

$$\bar{W} = E(W) = E(RI^2) = R\sigma_I^2, \quad \sigma_W^2 = E(R^2 I^4) - \bar{W}^2 = 2R^2\sigma_I^4$$

- Illustration: Table 2.1 and 2.2, Figure 2-14, Histogram
- Another Example: random number generation
  - Pp. 75-76 (randn.m), Figure 2-15 and 2-16
- More Illustration:
  - Exercise 2-5.1, 2-5.2

Some Gaussian-Derived Distributions

1. Distribution of Power (cf. p.8): *I* Gaussian $W = RI^2$

$$f_W(w) = \frac{1}{2\sqrt{Rw}}[f_I(\sqrt{w/R}) + f_I(-\sqrt{w/R})], \; w \geq 0$$

Show:

$$\bar{W} = E(W) = E(RI^2) = R\sigma_I^2, \;\; \sigma_W^2 = E(R^2 I^4) - \bar{W}^2 = 2R^2\sigma_I^4$$

More Illustrations:

- Figure 2-17 (pdf of *W*)
- Section 2-6, Exercise 2-6.1

Center of Signal and Image Processing
Georgia Institute of Technology
CSIP

More Gaussian-Derived Distributions

2. Rayleigh r.v.: X, Y independent Gaussian with same σ^2

$$R = \sqrt{X^2 + Y^2} \quad f_R(r) = \frac{r}{\sigma^2}\exp[-\frac{r^2}{2\sigma^2}],\ r \ge 0$$

$$\bar{R} = E(R) = \sqrt{\frac{\pi}{2}}\sigma,\quad \sigma_R^2 = E(R^2) - \bar{R}^2 = (2 - \frac{\pi}{2})\sigma^2$$

$$F_R(r) = \int_{-\infty}^{r} f_R(u)du = 1 - \exp[-\frac{r^2}{2\sigma^2}],\ r \ge 0$$

- More Illustrations:
  - Figure 2-18 (pdf of R), Exercise 2-6.2,
  - Maxwell, Chi-Square, Log-Normal r.v.'s, Exercises 2-6.3, 2-6.4, 2-6.5

Some Useful Distributions (III)

- *Uniform* Distribution: *U(X=x; a, b)*

$$U(x;a,b)=\begin{cases} 1/(b-a) & a\le x\le b \\ 0 & \text{otherwise} \end{cases} \quad \text{with } \ a<b$$

- Show: its mean, variance and distribution function

$$\mathrm{E}_U(X)=\frac{1}{2(b+a)} \quad Var_U(X)=\frac{1}{12}(b-a)^2$$

$$F_U(x)=\frac{x-a}{b-a}(\text{linear in } x),\ a\le x\le b$$

- More Illustration
  - Quantization error (Figure 2-20) and Phase distribution

Chi-Square Distributions

3. Chi-Square: sum of square iid N(0,1) random variables

$$X^2 = Y_1^2 + Y_2^2 + \cdots + Y_n^2 \text{ with } Y_1, \ldots, Y_n \text{ iid N(0,1) r.v.}$$

X^2 is said to be Chi-square with n degree of freedom: $\chi^2(n)$

$$f_{\chi^2}(u) = \frac{u^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\exp[-\frac{u}{2}],\ u \geq 0$$

Show that $\overline{U} = n$ and $\mathrm{Var}(U) = 2n$

- Implications: with proper normalization
  - Power random variable W is $\chi^2(1)$
  - Squared Rayleigh random variable R^2 is $\chi^2(2)$
  - Squared Maxwell random variable V^2 is $\chi^2(3)$

Generation of Random Numbers

- If X is uniform over $[0, 1]$, find a transformation function $q(x)$ so that $Y=q(X)$ is a r.v. with a desirable distribution $F_Y(y)$, i.e. generating random numbers of a desirable distribution from a uniform r.v. X:

$$F_Y(y) = P(Y \le y) = P(q(X) \le q(y)) = P(X \le x) = F_X(x)$$

$$y = F_Y^{-1}(F_X(x)) = F_Y^{-1}(x), \quad 0 \le x \le 1, \text{ i.e. } q(x) = F_Y^{-1}(x)$$

For Rayleigh $F_R^{-1}(r) = \sqrt{-8\ln[1-r]}$, so we want

$$Y = q(X) = \sqrt{-8\ln[1-X]}, \text{ with } X \text{ being } U(0,1)$$

- More Illustration
  - Example on p.90, Figure 2-21 and Exercise 2-7.1

Some Useful Distributions (IV)

- *Exponential* Distribution: E$(T=t;\tau)$

$$F_E(t)=1-\exp(-t/\tau),\ f_E(t)=\frac{1}{\tau}\exp(-t/\tau),\ t\geq 0$$

- Random variable of time interval between events
- Show its mean and variance: $\mathrm{E}_E(T)=\tau,$ and $\mathrm{Var}_{\mathrm{E}}(T)=\tau^2$
- Erlang Distribution: Random variable of the time interval between any event and the *k*th following event

$$f_k(t)=\frac{t^{k-1}}{\tau^k(k-1)!}\exp(-t/\tau),\ t\geq 0,\ k=1,2,3...,\ \mu_k=k\tau,\sigma_k^2=k\tau^2$$

  – Gamma distribution: make *k* continuous α and $\beta=1/\tau$
- More Illustration: Exercises 2-7.2 and 2-7.3

Conditional PDF's and pdf's

- *Conditional* Distribution (continuous variable case):

  $F(x|M) = P(X \le x|M) = P(X \le x, M)/P(M),\ P(M) > 0$
- Event $P(X \le x, M)$ is event ξ such that $X(\xi) \le x,$ and $\xi \in M$
- Properties: similar to those discussed in Slides 5 and 7
- M can be defined in many ways, e.g. $M = \{X \le m\}$
- More Illustration:
  - Section 2-8, pp. 98-101
  - Exercises 2-8.1 and 2-8.2
  - Figures 2-25 and 2-26

Engineering Applications

- Statistical Inference Problems
  - *Classification*: choose one of the stochastic sources
  - *Decision* and *Hypothesis Testing*: comparing two stochastic assumptions and decide on how to accept one of them
  - *Estimation*: given random samples from an assumed distribution, find "good" guess for the parameters
  - *Prediction*: from past samples, predict next set of samples
  - *Regression* (*Modeling*): fit a model to a given set of samples
- More useful distributions listed in Appendix B
- More Illustrations: Section 2-9
  - Examples 2-9.1 and 2-9.2

Information Theory & Shannon

- **Claude E. Shannon (1916-2001, from BL to MIT): Father of Information Theory, Modern Communication Theory …**
- **Entropy (Self-Information) – b*it*, amount of info in r.v.**

$$H(X) = -\sum_{x \in \mathrm{X}} p(x)\log_2 p(x) = \mathrm{E}[\log_2 \frac{1}{p(X)}] \quad 0\log_2 0 = 0$$

- **Study of English – Cryptography Theory, *Twenty Questions* game, Binary Tree and Entropy, etc.**
- **Concept of Code – Digital Communication, Switching and Digital Computation (optimal Boolean function realization with digital relays and switches)**
- **Channel Capacity – Source and Channel Encoding, Error-Free Transmission over Noisy Channel, etc.**
- **C. E. Shannon, "A Mathematical Theory of Communication", Parts 1 & 2, *Bell System Technical Journal*, 1948.**

Entropy of English

- **Markov Approximation to Probability of Letters**

$$P(L) = P(l_1)P(l_2 \mid l_1) \cdots P(l_{|L|} \mid l_1, \ldots, l_{|L|-1}) \quad k - gram$$

$$\approx P(l_1)P(l_2 \mid l_1) \cdots P(l_k \mid l_1, \ldots, l_{k-1}) \prod_{i=k+1}^{|L|} P(l_i \mid l_{i-1}, l_{i-2}, \ldots, l_k)$$

Model	Cross Entropy (bits)	Comments
Zeroth order	4.76	uniform letter log(27)
First order	4.03	unigram
Second order	2.8	bigram
Shannon's 2nd Experiment	1.34	human prediction

Summary

- **Today's Class**
  - Information about ECE3075, Summer 2004
  - Random variables
- **Reading Assignments**
  - Cooper & McGillem, Chapter 2
- **Class Next Week**
  - On Several Random Variables (Chapter 3)