Functions of Random Variables

- Example: let $X = U(0,1)$ and $Y = \frac{-1}{\lambda}\ln(1-X), \lambda > 0$

Find $F_Y(y)$

Functions of Random Variables

- Example: let $X = U(0,1)$ and $Y = \frac{-1}{\lambda}\ln(1-X), \lambda > 0$

Find $F_Y(y)$

$$0 < X < 1$$
$$-1 < -X < 0$$
$$0 < 1-X < 1$$
$$-\infty < \ln(1-X) < 0$$
$$0 < Y < \infty$$

$$F_Y(y) = P(Y \le y) = P(\frac{-1}{\lambda}\ln(1-X) \le y)$$

$$= P(1-X \ge e^{-\lambda y}) = P(X \le 1-e^{-\lambda y})$$

$$F_Y(y) = 1-e^{-\lambda y}$$
$$f_Y(y) = \lambda e^{-\lambda y}$$

Mean Variance

Properties of the mean:

$$1)E[cX] = cE[X]$$

$$2)E[c_1X_1 + c_2X_2] = c_1E[X_1] + c_2E[X_2]$$

Properties of the variance:

$$1)Var(X) \geq 0$$

$$2)Var(cX) = c^2Var(X)$$

$$3)Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Rules

- Exr. $1)X = c \Rightarrow Var(X) = 0$

  $2)Y = aX \Rightarrow Var(Y) = a^2Var(X)$

  $3)Y = X + b \Rightarrow Var(Y) = Var(X)$

  $4)Y = X_1 + X_2 \Rightarrow Var(Y) = Var(X_1) + Var(X_2)$

- Question Is

$$Var(X - Y) = Var(X) - Var(Y)$$

Example

Let X and Y be independent and

E(X)= 5, E(Y)=-3, and $\sigma_X = 2, \sigma_Y = 3$

Find the mean and he Std Deviation of

Z=3X-2Y-2

E[Z]= 3E[X]-2E[Y]-2 = 15+6-2=19

Var(Z)= 9Var(X)+4Var(Y) = 9.4+4.9 = 72

$\sigma_Z = \sqrt{72}$

Covariance

- The covariance is a measure of dependency between two variables
- Def. Cov(X,Y)= E[(X-E[X]) * (Y-E[Y])]

= E[X,Y]-E[X]*E[Y]

For dependent Variables X, Y

Var(X+Y)=Var(X)+Var(Y) + 2E[X,Y] – 2E[X].E[Y]

= Var(X)+Var(Y)+2Cov(X,Y)

Remark: Large Covariance → Large dependency

But If the covariance is 0, it does not mean they are independent.

Rules

(Prove)

1. Cov(X,X)=Var(X)
2. Cov(X,Y)=Cov(Y,X)
3. Cov. Is linear i.e.

Cov(aX+bY,Z)=aCov(X,Z)+bCov(Y,Z)

Exr. For X,Y independent U(0,1)

Find $Cov(X+2Y, X^2-Y)$

Correlation Analysis

- In probability theory and statistics, **correlation** (often measured as a correlation coefficient) indicates **the strength and direction of a linear relationship between two random variables**.
- **Correlation** refers to the departure of two variables from independence. In this broad sense there are several coefficients, measuring the degree of correlation, adapted to the nature of the data.

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y},$$

Correlation Analysis

- Remark: $-1 \le \rho_{(X,Y)} \le 1$
- The sign of ρ indicates the direction of the relationship;
  - ρ near 0 indicates no linear relationship,
  - ρ near 1 or -1 indicates a strong linear relationship.

Correlation Analysis

- Remark: $-1 \le \rho_{(X,Y)} \le 1$
- The sign of ρ indicates the direction of the relationship;
  - ρ near 0 indicates no linear relationship,
  - ρ near 1 or -1 indicates a strong linear relationship.
- Example: For X= U(0,1) find $\rho_{(X,X^2)}$

$Var(X) = \frac{1}{12}$ and $Var(X^2) = E[X^4] - (E[X^2])^2 = \frac{1}{5} - (\frac{1}{3})^2 = \frac{4}{45}$

$$\rho(X,X^2) = \frac{Cov(X,X^2)}{\sigma_X \sigma_{X^2}} = \frac{1/12}{\sqrt{1/12}\sqrt{4/45}} \approx 9.96$$

i.e. very strong dependency

Joint Probability

- **Joint Probability (A ∩ B)**
  - The probability of two events in conjunction. It is the probability of both events together.

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

- **Independent Events**
  - Two events A and B are **independent** if

$$p(A \cap B) = p(A) \cdot p(B)$$

Joint Distribution

- For two random variables X and Y

$$F_{(X,Y)}(x,y) = p\{X \le x, Y \le y\}$$

- Example: For X,Y ind. U(0,1) r.v.

$$F_{(X,Y)}(x,y) = P(X \le x, Y \le y) = P(X \le x).P(Y \le y)$$

$$= F_X(x).F_Y(y) = x.y$$

$$f_{(X,Y)}(x,y) = \frac{\partial^2 F_{(X,Y)}(x,y)}{\partial x \partial y} = 1$$

Joint Probability density function

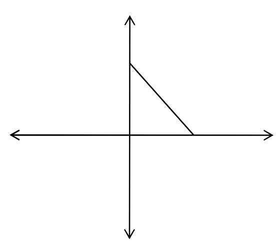
- X,Y are called jointly continuous if there exists f(x,y)
- S.t. $P(X \in A, y \in B) = \int_B \int_A f(xy)dxdy$
- If X, Y are independent then

$$f(x,y) = f_X(x) f_Y(y)$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

Example



- Find the Cor(X,Y) for

$$f(x,y) = \begin{cases} 24xy & x \ge 0, y \ge 0 \quad x+y \le 1 \\ o & o.w. \end{cases}$$

Find $Cov(X,Y) = E(XY) - E(X)E(Y)$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dydx = \int_0^1 \int_0^{1-x} xyf(x,y)dydx$$

$$= \int_0^1 \int_0^1 xy(24xy)dydx = \int_0^1 24x^2 \int_0^{1-x} y^2 dydx$$

$$= \int_0^1 24x^2 \frac{1}{3} y^3 |_0^{1-x} \; dx = \int_0^1 8x^2(1-x)^3 dx = \frac{2}{15}$$

Computing the covariance

$$E[X]=\int_1^1 x f_x(x)dx$$

$$f_x(x)=\int_1^{1-x} 24xydy=12xy^2 |_0^{1-x}=12x(1-x)^2$$

$$f_y(y)=\int_1^{1-y} 24xydx=12x^2y |_0^{1-y}=12y(1-y)^2$$

$$E[x]=\int_0^1 12x(1-x)^2dx=\frac{2}{5}$$

$$E[y]=\int_0^1 12y(1-y)^2dy=\frac{2}{5}$$

$$Cov(X,Y)=E[XY]-E[X]E[Y]=\frac{2}{15}-\frac{2}{5}.\frac{2}{5}=\frac{-2}{75}$$

Computing the Correlation

$$E[X^2]=\int_0^1 x^2 f_x(x)dx=\int_0^1 x^2 12x(1-x)^2dx$$

$$=\int_0^1 12x^3(1-2x+x^2)dx=4x^4-\frac{24}{5}x^5+\frac{12}{6}x^6 |_0^1$$

$$=3-\frac{24}{5}+\frac{12}{6}$$

$$Var(X)=E[X^2]-(E[X])^2=3-\frac{24}{5}+\frac{12}{6}-\frac{4}{25}=\frac{1}{25}$$

$$\text{Similarly}\, Var(X)=\frac{1}{25}$$

$$Cor(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{-2/75}{1/25}=\frac{-2}{3}$$

Conditional Probability

- **Joint Probability (A ∩ B)**: the probability of two events in conjunction. That is, it is the probability of both events together. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
- **Independent Events:** Two events A and B are **independent** if $p(A \cap B) = p(A) \cdot p(B)$

- **Conditional Probability p(A|B)** is the probability of some event A, **given** B. $$p(A \mid B) = \frac{p(A \cap B)}{p(B)}$$
- p(A) is the prior probability; p(A|B) is called a posterior probability.

Remark: If A and *B* are **independent**,
then $p(A \mid B) = p(A)$ and $p(B \mid A) = p(B)$

Baye's Rule

- since $p(A \cap B) = p(B \cap A)$
- Using Conditional Probability definition, we have
$$p(A \mid B) \cdot P(B) = p(B \mid A) \cdot P(A)$$
- The Bayes rule is:
$$p(A \mid B) = \frac{p(B \mid A) \cdot p(A)}{p(B)}$$

Law of Total Probability

- In probability theory, the law of total probability means the **prior probability** of A, P(A), is equal to the expected value of the **posterior probability** of A.
- That is, for any random variable B,

$$P(A) = E[P(A/B)]$$

  - where $p(A|B)$ is the conditional probability of A given B.

Law of Total Probability

- **Law of alternatives:** The term ***law of total probability*** is sometimes taken to mean the **law of alternatives**, which is a special case of the law of total probability applying to discrete random variables.
- if $\{ B_n : n = 1, 2, 3, \ldots \}$ is a ***finite*** partition of a probability space and each set B_n is measurable, then for any event A we have

$$p(A) = \sum_n p(A \cap B_n)$$

or, alternatively, (using Rule of Conditional Probability)

$$p(A) = \sum_n p(A \mid B_n) \cdot p(B)$$

Ex Law of Total Probability

Sample Space $S = \{1,2,3,4,5,6,7\}$

Partitions $B1 = \{1,5\}$ $B2 = \{2,3,6\}$ $B3 = \{4,7\}$

Event $A = \{3\}$

Law of Total Probability

$$p(A) = p(\{3\}) = p(A \cap B_1) + p(A \cap B_2) + p(A \cap B_3)$$
$$= 0 + p(\{3\} \cap \{2,3,6\}) + 0 = p(\{3\})$$

Random Number Generator Chapter 7

- In simulations, we generate random values for variables with a specified distribution
- Ex., model service times using the exponential distribution
- Generation of random values is a two step process

  Random number generation: Generate random numbers uniformly distributed between 0 and 1

  Random variate generation: Transform the random numbers generated above to obtain numbers satisfying the desired distribution

Sources of Randomness for Common Simulation Applications

Type of system	Sources of randomness
Manufacturing	Processing times, machine times to failure, machine repair times
Defense-related	Arrival times and payloads of missiles or airplanes, outcome of an engagement
Communications	Interarrival times of messages, message types, message lengths
Transportation	Ship-loading times, interarrival times of customers to a subway

23

Random Numbers

- True Random Numbers
- Pseudo random numbers
- Quasi random numbers

True Random Numbers

- The most desirable random numbers
- Generated only from physical experiment
- EX: - Ten sided die
  - Throwing a coin assuming 1 for head (0 for tail)
  - Recording the times the customers arrive
  - Physical devices “ Geiger counter”, recording the interval between the decay of a radioactive material

Disadvantages of True Random Numbers

- Slow
- Inconvenient
- Expensive to build
- Not reproducible

Pseudo Random Numbers

- Algorithms can automatically create long runs (for example, millions of numbers long) of random numbers with good random properties but eventually the sequence repeats.
- determine the next random number as a function of the previously generated random number (i.e., recursive calculations are applied)
- Random numbers generated, are therefore, deterministic. That is, sequence of random numbers is known given the starting seed.
- Random number generators have a cycle length (or period) that is measured by the count of unique numbers generated before the cycle repeats itself.

Pseudo Random Numbers

- Most computer programming languages include functions or library routines that surport random number generators.
- Such library functions often have poor statistical properties and some will repeat patterns after only tens of thousands of trials. These functions may provide enough randomness for certain tasks but are unsuitable where high-quality randomness is required, such as in cryptographic applications, statistics or numerical analysis.

Advantages of Pseudo Random Numbers

- Uniformly distributed between 0, 1
- Satisfiably independent: produce output satisfying statistical tests of randomness
- Reproducible: ability to reproduce random number stream if necessary
- Not repeated for a long desired length
- Generated fast
- Does not need big memory to store: easy to move random number generator to a new machine

❖ Form clusters and empty regions in higher dimensions

Quasi Random Numbers

- Also Called low discrepancy sequences
- Discrepancy of a sequence is a measure of its uniformity
- Computed by comparing the actual numbers of points in multidimensional space to the number assuming full uniform distribution
- Not random at all.
- Yet, Uniformly fill the space

Why is this Important?

- Validity
  - The simulation model may not be valid due to cycles and dependencies in the model
- Precision
  - You can improve the output analysis by carefully choosing the random numbers

Von Neuman Midsquare method 1940

- Let m = the number of digits requires after the decimal pt
- Start with a seed
- Square the seed
- Consider only the m middle digits
- EX: m = 2 , seed = 23
- Halt when you get 0 or

equal numbers

Generated numbers

.23 , .52,.70,.90,.10

0 [52] 9
2 [70] 4
4 [90] 0
8 [10] 0
0 [10] 0

Linear Congruential Generator

- Producing a sequence of integers, x_0, x_1, $x_{2,}$ … between *0 and m-1 by following a recursive* relationship:

$$X_{i+1}= (a X_i + c) \bmod m \qquad i =0,1,2,...$$

- The selection of the values for *a, c, m, and* X_0 affects the statistical properties and the cycle length.
- The random integers are being generated *[0,m-1],* and to convert the integers to random numbers:
- $R_i= X_i/m$, i =1,2,…
- Halt when a number is repeated

LCG Ex.

- Use X_0 *= 27, a = 17, c = 43, and m = 100.*

$$X_{i+1}= (a X_i + c) \bmod m$$

The X_i *and* R_i *values are:*

- X_1 *= (17*27+43) mod 100*
  = 502 mod 100 = 2, R_1 *= 0.02;*
- X_2 *= (17*2+32) mod 100 = 77,* R_2 *= 0.77;*
- X_3 *= (17*77+32) mod 100 = 52,* R_3 *= 0.52;*
- …

LCG

Remarks:

- c = 0, the generator is called Multiplicative LCG.
- If c ≠ 0, the generator is called Mixed LCG.
- The length of the cycle is called its Period, can be at most?
- m should be chosen to be big
- Choose m of the form 2^k for efficient computation

LCG

Remarks:

- c = 0, the generator is called Multiplicative LCG.
- If c ≠ 0, the generator is called Mixed LCG.
- The length of the cycle is called its Period, can be at most m-1
- m should be chosen to be big
- Choose m of the form 2^k for efficient computation

LCG

Theorem:

If c ≠ 0, LCG has full period

iff

- Integers *m* and *c* are relatively prime (the only positive integer that divides both *m* and c is 1)
- Every prime number that is a factor of *m* is also a factor of a-1
- If integer *m* is a multiple of 4, a-1 is also a multiple of 4

Multiplicative Congruential Generator

- If c= 0, X_{i+1}= (a X_i) mod m i =0,1,2,…
- Break the first condition of the thm. i.e. Not full period
- Its max period is 2^{b-1} where b is the # of digits (size of ward)
- Advantages:

Faster, Simpler, easy to implement