Thm 3.13

Linear Transformation of a Gaussian gives another Gaussian.

$$Y = aX + b \quad \text{where} \quad X \sim \text{Gaussian}(\mu, \sigma)$$

$$P(Y \le y) = P[aX + b \le y]$$

$$= P\left[X \le (y-b)/a\right]$$

$$= \int_{-\infty}^{(y-b)/a} f_X(x)\, dx$$

$$= \int_{-\infty}^{(y-b)/a} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}\, dx \quad \text{— (1)}$$

$$y = \frac{a(y-b)}{a} + b$$

let $ax + b = z$

$\therefore \quad x = \left(\frac{z-b}{a}\right) \Rightarrow z = ax + b.$

$$dx = \frac{dz}{a}$$

$\therefore$ (1) becomes

$$\int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{z-b}{a} - \mu\right)^2/2\sigma^2}\, dz$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi(a\sigma)^2}} e^{-\left(z - (a\mu + b)\right)/[2(a\sigma)^2]}\, dz.$$

$$\Rightarrow \quad Y \sim \text{Gaussian}(a\mu + b,\ a\sigma)$$

Hence proved.

$$Y = aX + b$$

$$f_Y(y) = f_Y(ax+b)$$

$$F_Y(y) = F_Y(ax+b)$$

$$f_Y(y)\underset{dy}{\downarrow} = f_{aX+b}(y)\,dy = f_X\left(\frac{y-b}{a}\right)\frac{dy}{a}$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\left(\frac{\left(\frac{y-b}{a} - \mu_X\right)^2}{2\sigma_X^2}\right)} \frac{dy}{a}$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2 a^2}} e^{-\left(\frac{(y-b-a\mu_X)^2}{2\sigma_X^2 a^2}\right)} dy.$$

$X = 0 \Rightarrow Y = b.$ There is a one to one correspondence. We want rate of acc around $X=0$ to be same as around $Y=b$.

Probability and Random Processes

Continuous Random Variables

Sanjit Kaul

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Continuous Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Intervals

- (x_1, x_2)
  - Includes neither x_1 nor x_2 but all real numbers in between
- $[x_1, x_2)$
- $(x_1, x_2]$
- $[x_1, x_2]$

A Very Narrow Carrom Board

- A very narrow board of length L can be approximated by a line
  - This is a terrible attempt on my part to emulate the example in RY
- You shoot the striker (red)
- Your strike is powerful
- This ensures that the striker stops at any point along the thin board with equal probability
- The striker stops at a point in [0,L]

A Very Narrow Carrom Board

- Let X be the RV that describes the final location of the pointer
- We have $S_X = [0,L]$
- The range is a continuum. Thus X is a CRV
- What is $P[X = x]$ where $x \in S_x$?

A Very Narrow Carrom Board

- Let's think in discrete terms

- We split the board into n equal intervals
- Define the RV Y with $S_Y = \{1,2,\ldots,n\}$
- Y is a discrete uniform RV
- Note that in the discrete problem all values of X in interval k correspond to Y=k.

$$P[Y = k] = P[k - 1 \leq X < k]$$

A Very Narrow Carrom Board

$$\text{If } X = x \text{ then } Y = \lceil x/(L/n) \rceil = \lceil nx/L \rceil$$

- Also

$$\{X = x\} \subset \{Y = \lceil nx/L \rceil\}$$

- Therefore

$$P[X = x] \leq P[Y = \lceil nx/L \rceil] = 1/n$$

A Very Narrow Carrom Board

- Getting as close as possible to the continuous case

$$P[X = x] \leq \lim_{n \to \infty} 1/n = 0$$

- Therefore

$$P[X = x] = 0$$

- Every outcome {X=x} has a probability 0!
  - Note that this is not true for an interval on the board
- The PMF is not a useful way of describing a continuous RV
- What about the CDF?

Def 3.1 Cumulative Distribution Function

- The CDF of a RV X is $F_X(x) = P[X \le x]$
- Properties of a CDF are the same for all kinds of random variables
  - Discrete,
  - Continuous, and
  - Mixed.

Theorem 3.1

For any random variable X,

(a) $F_X(-\infty) = 0$

(b) $F_X(\infty) = 1$

(c) $P[x_1 < X \le x_2] = F_X(x_2) - F_X(x_1)$

$P[X \in (x_1, x_2]]$

Problem

- What is the CDF of X?
- $P[X \leq x] = ?$, What is your guess?
- We can say that
  - $P[X \leq x] = 1$, for $x >= L$
  - $P[X \leq x] = 0$, for $x <= 0$
- The probability accumulates and goes from 0 to 1 as we go from x=0 to x=L.
- Given that all regions on the board are equally preferred, intuitively probability must accumulate from 0 to x/L as we go from 0 to x
  - $P[X \leq x] = x/L$ for $0 <= x <= L$

Problem

- We have

$$\{Y \leq \lceil nx/L \rceil - 1\} \subset \{X \leq x\} \subset \{Y \leq \lceil nx/L \rceil\}$$

- Therefore

$$F_Y(\lceil nx/L \rceil - 1) \leq P[X \leq x] \leq F_Y(\lceil nx/L \rceil)$$

- Since Y is a discrete uniform distribution

$$\frac{\lceil nx/L \rceil - 1}{n} \leq P[X \leq x] \leq \frac{\lceil nx/L \rceil)}{n}$$

Calculating CDF for the Board Game

- As n goes to infinity, we get the CDF of the continuous uniform RV

$$P[X \leq x] = x/L \text{ for } 0 \leq x < L.$$

- We used the fact that

$$\lim_{n\to\infty} \frac{\lceil nx/L \rceil}{n} = \frac{x}{L}$$

CDF Discrete vs. Continuous

- The CDF of a discrete RV contained jumps
- The size of the jump (change) was the probability of the RV at the point of jump
- The larger the jump, the larger the probability
- The CDF of a continuous RV is continuous

Quiz 3.1

$f(x) = x^2$

The cumulative distribution function of the random variable Y is

$(-\infty, y]$

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y/4 & 0 \le y \le 4 \\ 1 & y > 4 \end{cases} \tag{3.9}$$

Sketch the CDF of Y and calculate the following probabilities:

$f(x=1)$

$f(1)$

(1) $P[Y \le -1]$ $\quad F_Y(y=1) - F_Y(-\infty)$

(2) $P[Y \le 1]$ $\quad F_Y(y \le 1)$

(3) $P[2 < Y \le 3]$ $\quad F_Y(3) - F_Y(2)$ $\quad P[Y \le 3] - P[Y \le 2]$

(4) $P[Y > 1.5]$

Example CDF of a continuous RV

Figure 3.2

$F_X(x_3+\Delta) - F_X(x_3)$

$\leq$

The graph of an arbitrary CDF $F_X(x)$.

- From the figure

$$F_X(x_2+\Delta) - F_X(x_2) > F_X(x_1+\Delta) - F_X(x_1)$$

Information in a CDF

Figure 3.2

$$\lim_{\Delta \to 0} \frac{F_X(x_2+\Delta) - F_X(x_2)}{(x_2+\Delta) - x_2}$$

$f_X(x_2)$ PDF

The graph of an arbitrary CDF $F_X(x)$.

- Rate of accumulation of probability in the region right of x_2 is greater than the in the region right of x_1
- As the interval Δ becomes smaller the rate of accumulation -> slope of the CDF

Probability Density Function

- Greater **the slope of the CDF** at any point x, the more the likelihood of an outcome in an interval around x

Probability Density Function

Definition 3.3 *(PDF)*

The probability density function *(PDF) of a continuous random variable* X *is*

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

Note the lowercase f

Example PDF

Figure 3.3

The PDF of the modem receiver voltage X.

- You are sending 0(s) and 1(s) mapped to -5 and +5
- Receiver sees the above PDF, thanks to Gaussian noise
- Which regions are most likely to be seen by the receiver and why?

Properties of a PDF

Theorem 3.2

For a continuous random variable X with PDF $f_X(x)$,

(a) $f_X(x) \geq 0$ for all x,

(b) $F_X(x) = \int_{-\infty}^{x} f_X(u)\,du$,

(c) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$.

- Why is (a) true?
  - CDF is a non-decreasing function and PDF is its slope
- Why (b)?
  - From definition of a PDF
- Why (c)?
  - The integral is $P[X < \infty] = F_X(\infty) = 1$.

Properties of a PDF

Theorem 3.3

$$P[x_1 < X \le x_2] = \int_{x_1}^{x_2} f_X(x)\, dx.$$

$$P[X \le x_2] - P[X \le x_1]$$

$$\int_{-\infty}^{x_2} f_X(x)\, dx - \int_{-\infty}^{x_1} f_X(x)\, dx$$

The PDF and CDF of X.

Properties of a PDF

Figure 3.2

The graph of an arbitrary CDF $F_X(x)$.

$$P[x_2 < X \leq x_2 + \Delta] = \frac{F_X(x_2+\Delta) - F_X(x_2)}{\Delta}\Delta$$

Figure 3.2

The graph of an arbitrary CDF $F_X(x)$.

$$P[x_2 < X \leq x_2 + \Delta] = \frac{F_X(x_2+\Delta)-F_X(x_2)}{\Delta}\Delta$$

- As the interval Δ becomes smaller

$$\frac{F_X(x_2+\Delta)-F_X(x_2)}{\Delta} = f_X(x_2)$$

- In general we can write for any x and an infinitesimally small *dx*

$$P[x < X \leq x + dx] = f_X(x)dx$$

Intervals and CRV(s)

- Consider the four different events
- A = (0,1)
- B = (0,1]
- C = [0,1)
- D = [0,1]
- They belong to the range space of a **continuous RV X**
- What can we say about P[A], P[B], P[C], and P[D]?
- Since X is continuous, P[A] = P[B] = P[C] = P[D]
  - This is true because probability of a point (outcome) is 0 for a continuous RV

Problem on PDF

Quiz 3.2

Random variable X has probability density function

$$f_X(x) = \begin{cases} cxe^{-x/2} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.25}$$

Sketch the PDF and find the following:

(1) the constant c

(2) the CDF $F_X(x)$

(3) $P[0 \leq X \leq 4]$

(4) $P[-2 \leq X \leq 2]$

(5) What is the mode of the random variable X?

Problem 3.1.3

$P[W \le 3]$

The CDF of random variable W is

$y = (x+5)/8$

$$F_W(w) = \begin{cases} 0 & w < -5, \\ (w+5)/8 & -5 \le w < -3, \\ 1/4 & -3 \le w < 3, \\ 1/4 + 3(w-3)/8 & 3 \le w < 5, \\ 1 & w \ge 5. \end{cases}$$

$P[W \le w]$

(a) What is $P[W \le 4]$?

(b) What is $P[-2 < W \le 2]$?

(c) What is $P[W > 0]$?

(d) What is the value of a such that $P[W \le a] = 1/2$?

$\frac{1}{4} + 3(w-3)/8 = 1/2$

$w = 11/3$

$P[W \le 2.9] = 1/4$

Problem 3.2.4

For a constant parameter $a > 0$, a Rayleigh random variable X has PDF

$$f_X(x) = \begin{cases} a^2 x e^{-a^2 x^2/2} & x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

What is the CDF of X?

Expected Value

- **Def 3.4** The expected value of a continuous random variable X is

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

- Summation for the discrete case is replaced by integration for the continuous case
- **Theorem 3.4** The expected value of a function g(X) of a continuous random variable X is

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

Example of g(X)

$$Var[X] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx$$

$P[X \leq x] = x/L$ for $0 \leq x < L$. (and 0 otherwise)

What is $f_X(x)$? $\Rightarrow$ $f_X(x) = \begin{cases} \frac{1}{L} & 0 \leq x < L \\ 0 & \text{otherwise.} \end{cases}$

VVVS Problem

- What is $E[X]$?
  - $L/2$
- What is $E[X^2]$?
  - $L^2/3$
- What is $Var[X]$?
  - $L^2/3 - L^2/4 = L^2/12$

These Equalities are Valid for Both Continuous and Discrete RV(s)

Theorem 3.5

For any random variable X,

(a) $E[X - \mu_X] =$

(b) $E[aX + b] =$

(c) $\mathrm{Var}[X] =$

(d) $\mathrm{Var}[aX + b] =$

Problem 3.4.1

Radars detect flying objects by measuring the power reflected from them. The reflected power of an aircraft can be modeled as a random variable Y with PDF

$Y > E[Y]$

$$f_Y(y) = \begin{cases} \frac{1}{P_0} e^{-y/P_0} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $P_0 > 0$ is some constant. The aircraft is correctly identified by the radar if the reflected power of the aircraft is larger than its average value. What is the probability $P[C]$ that an aircraft is correctly identified?

$\frac{1}{P_0}\int_0^\infty e^{-y/P_0}\,dy = 1$ $\quad \int_0^\infty y\, f_Y(y)\,dy$ $\quad P_0$ $\quad e^{-1}$ $\quad \int_{P_0}^\infty f_Y(y)\,dy$

Families of Continuous RV(s)

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Uniform(a,b)

Definition 3.5 Uniform Random Variable

X *is a uniform* (a, b) *random variable if the PDF of* X *is*

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

where the two parameters are $b > a$.

- The probability of an outcome being in any interval of a given size Δ is the same
- How can we conclude this?
- The final position of the striker in the board example was *uniformly distributed*

Theorem 3.7

Let X be a uniform (a, b) random variable, where a and b are both integers. Let $K = \lceil X \rceil$. Then K is a _____ random variable.

- You want to describe the RV K
- K is a discrete RV. It is described completely by its PMF P[K = k]

$P[K=k]$

Given the ceil function, all x that are > k-1 and <= k correspond to the outcome {K=k}

$$P[K = k] = P[k - 1 < x \leq k] = \frac{1}{b-a}$$

Discrete and Continuous Uniform RV(s)

$$P[K = k] = P[k - 1 < x \leq k] = \frac{1}{b-a}$$

- The above is true for k = a+1, a+2,…, b
- P[K=k] = 0 otherwise.
- Clearly K is a discrete uniform RV

Exponential(λ)

Definition 3.6 *Exponential Random Variable*

X *is an* exponential (λ) random variable *if the PDF of* X *is*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & \textit{otherwise}, \end{cases}$$

where the parameter $\lambda > 0$.

A

$\cancel{f_X(x)}$

$$f_Y(y) = \begin{cases} c\lambda e^{-\lambda y} & 0 \leq y < a \\ 0 & \text{otherwise} \end{cases}$$

- The inter-arrival time between two packet arrivals at a server may be modeled as an Exp RV
- What is the average inter-arrival time?

$c = \frac{1}{A}$

Exponential and Geometric RV(s)

Theorem 3.9

If X is an exponential (λ) random variable, then $K = \lceil X \rceil$ is a geometric (p) random variable with $p = 1 - e^{-\lambda}$.

- Proof is similar to the case of Uniform distribution

Example 3.14 Problem

Phone company A charges $\$0.15$ per minute for telephone calls. For any fraction of a minute at the end of a call, they charge for a full minute. Phone Company B also charges $\$0.15$ per minute. However, Phone Company B calculates its charge based on the exact duration of a call. If T, the duration of a call in minutes, is an exponential $(\lambda = 1/3)$ random variable, what are the expected revenues per call $E[R_A]$ and $E[R_B]$ for companies A and B?

Erlang(n,λ)

Definition 3.7 ***Erlang Random Variable***

X *is an* Erlang (n, λ) *random variable if the PDF of* X *is*

$$f_X(x) = \begin{cases} \dfrac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

where the parameter $\lambda > 0$*, and the parameter* $n \geq 1$ *is an integer.*

- The waiting time between n-occurrences of an event
- The waiting time between n customer arrivals
- n => Sum of n independent Exp(λ) Random Variables

Theorem 3.10

If X is an Erlang (n, λ) random variable, then

$$X = X_1 + X_2 + \cdots + X_n$$

$$E[X] = \frac{n}{\lambda}, \qquad \mathrm{Var}[X] = \frac{n}{\lambda^2}.$$

$X_1, X_2, \ldots, X_n$

$X_i \sim f_{X_i}(x) \sim \mathrm{Exp}(\lambda)$

$E[X_i] = \frac{1}{\lambda}$

$E[X]$

$= E[X_1] + \cdots + E[X_n]$

$= \frac{n}{\lambda}$

$E[(X_1 + \cdots + X_n - E[X_1 + \cdots + X_n])^2]$

$= E[(X_1 + X_2 + \cdots + X_n)^2] - (E[(X_1 + \cdots + X_n)])^2$

Quiz 3.4

Continuous random variable X has $E[X] = 3$ and $\mathrm{Var}[X] = 9$. Find the PDF, $f_X(x)$, if

(1) X has an exponential PDF,

(2) X has a uniform PDF.

(a, b)

$$\frac{a+b}{2} = 3$$

$$\frac{(b-a)^2}{12} = 9$$

A, A^c $\quad A = \{M \le 20\}$ $\quad E[C_B] = E[C_B|A]P[A]$

Problem 3.4.9

$A^c = \{M > 20\}$ $\quad + \underbrace{E[C_B|A^c]}_{E[99+10(M-20)]} P[A^c]$

Long-distance calling plan A offers flat rate service at 10 cents per minute. Calling plan B charges 99 cents for every call under 20 minutes; for calls over 20 minutes, the charge is 99 cents for the first 20 minutes plus 10 cents for every additional minute. (Note that these plans measure your call duration exactly, without rounding to the next minute or even second.) If your long-distance calls have exponential distribution with expected value τ minutes, which plan offers a lower expected cost per call?

M

$C_A = 10M$ $\quad E[C_A] = 10E[M] = 10\tau.$

$C_B = 99 + 10(M-20)$ $\quad M \ge 20$

$P[M \le 20]$

$P[M > 20]$

$E[C_B] = P[M \le 20]\,99 + P[M > 20]\,E[99 + 10(M-20)]$

$99+$

$$P[M>20] = e^{-(\frac{1}{\tau})20}$$

$$\underbrace{(M-20)}_{\hookrightarrow}10$$

$$M \sim Exp\left(\frac{1}{\tau}\right)$$

$$P[M>m] = e^{-M/\tau}$$

$$P[M-20>10] = P[M>30]$$

$$E[M-20] = E[M'] = \int_0^\infty m'$$

$$f_{M|M>20}(m)$$

$$C_B = 99 + (X-20)^+ 10$$

$$E[C_B] = E[g(X)]$$

$$= E\left[99 + (X-20)^+ 10\right]$$

$$\int_0^\infty g(x) f_X(x)\,dx$$

$$(x-20)^+$$

$$= 99 + 10E\left[(X-20)^+\right]$$

$$= \int_{20}^\infty (x-20) f_X(x)\,dx$$

Gaussian(μ,σ)

- The Bell-shaped distribution
- The Normal distribution
- Parameters are the mean μ and standard deviation is σ
  - Variance is σ^2
  - If X is Gaussian we often write X is N[μ, σ^2]

Definition 3.8 ***Gaussian Random Variable***

X is a Gaussian (μ, σ) random variable if the PDF of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$

where the parameter μ can be any real number and the parameter $\sigma > 0$.

Figure 3.5

(a) $\mu = 2, \sigma = 1/2$ **(b)** $\mu = 2, \sigma = 2$

Two examples of a Gaussian random variable X with expected value μ and standard deviation σ.

Show that the Gaussian PDF is a valid PDF

- Let W be a Gaussian RV

$$\int_{-\infty}^{\infty} f_W(w)\,dw = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(w-\mu)^2/(2\sigma^2)}\,dw$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(x)^2/(2\sigma^2)}\,dx$$

- The square of the integral is

$$I^2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(y)^2/(2\sigma^2)}\,dy \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(x)^2/(2\sigma^2)}\,dx$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} e^{(x^2+y^2)/(2\sigma^2)}\,dx\,dy$$

- Substitute x=r sin(ϕ) and y = rcos(ϕ) and show $I^2 = 1$

Derive $E[W]$ and $E[W^2]$

- Use steps similar to showing area is unity
- You will also need to apply integration by parts

Linear Transformation of a Gaussian

Theorem 3.13

If X is Gaussian (μ, σ), $Y = aX + b$ is _____.

$a^2\sigma^2$

- Linear transformation of a Gaussian gives another Gaussian!
- How do show the above?
- If X is not Gaussian, what are E[Y] and E[Y^2]?

Y is Gaussian $(a\mu+b, a\sigma)$

Standard Normal Variable and CDF

- **Def 3.9** The standard random variable is Gaussian(0,1) – 0 mean and unit variance

(μ, σ)

$$\int_{-\infty}^{x} f_X(x)\,dx$$

$$\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-[(x-\mu)^2/(2\sigma^2)]}$$

Definition 3.10 ***Standard Normal CDF***

The CDF of the standard normal random variable Z is

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\,du.$$

$$E(\sigma Z + \mu)$$

$X \sim \text{Gaussian}(\mu, \sigma)$ $\quad Z \sim N(0,1)$ $\quad \frac{X-\mu}{\sigma} = Z$

Expressing a Gaussian CDF as a N[0,1] CDF

Theorem 3.14

$P[X \le 100] = P[X - \mu_X \le 100 - \mu_X]$

$= P\left[\frac{X-\mu_X}{\sigma_X} \le \frac{100-\mu_X}{\sigma_X}\right]$

If X is a Gaussian (μ, σ) random variable, the CDF of X is

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

$X \sim N(2, 20)$

The probability that X is in the interval $(a, b]$ is

$P[X \le 100]$

$$P[a < X \le b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

$= \int_{-\infty}^{100} f_X(x)\, dx$

$\int_{-\infty}^{100} \frac{1}{\sqrt{2\pi(400)}} e^{-\left[(x-2)^2/2(400)\right]} dx$

- Show using definition of CDF and substitution of variables in the integral
- All I need are values of ϕ(.)
  - This is important as CDF integral calculations for finite limits can only be done numerically
  - Thankfully, all we now need is a tabulation of ϕ(.)

Standard Normal CDF

- In fact $\phi(-z) = 1 - \phi(z)$
  - Why?
  - Is this true for any Gaussian Distribution?
  - When is it true?
- All I need is $\phi(z)$ for $z >= 0$

Figure 3.6

$$\Phi(-z) = \int_{-\infty}^{-z} f_Z(z)\,dz$$

$$= \int_{-\infty}^{-z} \frac{1}{\sqrt{2\pi(1)}} e^{-((z-0)^2/(2(1)))}\,dz$$

$$\int_{+\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\,(-dy) = \int_{-\infty}^{-z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$

$$1-\Phi(z) = \int_{-\infty}^{\infty} f_Z(z) - \int_{-\infty}^{z} f_Z(z)\,dz$$

$$\int_{z}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$

μ
z

Example 3.16 Problem

If X is the Gaussian $(61, 10)$ random variable, what is $P[X \leq 46]$?

$$F_X(46) = \Phi\left(\frac{46-61}{10}\right)$$
$$= \Phi(-1.5) = 1 - \Phi(1.5)$$

$$P[X \leq 46] = P[X - \mu_X \leq 46 - \mu_X]$$
$$= P\left[\frac{X-\mu_X}{\sigma_X} \leq \frac{46-\mu_X}{\sigma_X}\right]$$
$$= \Phi\left(\frac{46-\mu_X}{\sigma_X}\right)$$

$Z \sim N(0, 1)$

$\mu_Z \quad \sigma_Z$

$X \sim N(\mu_X, \sigma_X)$

$$\frac{X - \mu_X}{\sigma_X} \rightarrow Z$$

Example 3.17 Problem

If X is a Gaussian random variable with $\mu = 61$ and $\sigma = 10$, what is $P[51 < X \leq 71]$?

$$P[51 < X \leq 71] = P[X \leq 71] - P[X \leq 51]$$

$$\frac{X - \mu}{\sigma_X}$$

$$\frac{X - \mu}{\sigma_X^2}$$

Standard Normal CCDF

Definition 3.11

CCDF

The standard normal complementary CDF *is*

$$Q(z) = P[Z > z] = \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-u^2/2}\, du = 1 - \Phi(z).$$

$$\Phi(3) = 0.9987, \;\; \Phi(4) = 0.9999768$$

$$Q(3) = 1.35 \times 10^{-3}, Q(4) = 3.17 \times 10^{-5}$$

Problem 3.5.6

A professor pays 25 cents for each blackboard error made in lecture to the student who points out the error. In a career of n years filled with blackboard errors, the total amount in dollars paid can be approximated by a Gaussian random variable Y_n with expected value $40n$ and variance $100n$. What is the probability that Y_{20} exceeds 1000? How many years n must the professor teach in order that $P[Y_n > 1000] > 0.99$?

$$P[Y_{20} > 1000] = P\left[\frac{Y_{20} - 40(20)}{10\sqrt{20}} > \frac{1000 - 40(20)}{10\sqrt{20}}\right]$$

$$P\left[\frac{Y_n - 40n}{10\sqrt{n}} > \frac{1000 - 40n}{10\sqrt{n}}\right] > 0.99$$

Problem 3.5.7

Suppose that out of 100 million men in the United States, 23,000 are at least 7 feet tall. Suppose that the heights of U.S. men are independent Gaussian random variables with a expected value of $5'10''$. Let N equal the number of men who are at least $7'6''$ tall.

(a) Calculate σ_X, the standard deviation of the height of men in the United States.

(b) In terms of the $\Phi(\cdot)$ function, what is the probability that a randomly chosen man is at least 8 feet tall?

(c) What is the probability that there is no man alive in the U.S. today that is at least $7'6''$ tall?

(d) What is $E[N]$?

$$P[X > 7'] = \frac{23000}{10^8}$$

$$P\left[\frac{X - 5'10''}{\sigma_X} > \frac{7' - 5'10''}{\sigma_X}\right] = 23 \times 10^{-5}$$

- Six Sigma Event?
  - Use MATLAB's qfunc

$$(\mu - 6\sigma, \mu + 6\sigma)$$

$$(\mu - \sigma, \mu + \sigma)$$

$$(\mu - 2\sigma, \mu + 2\sigma)$$

Quiz 3.5

X is the Gaussian $(0, 1)$ random variable and Y is the Gaussian $(0, 2)$ random variable.

$(\mu - \sigma, \mu + \sigma)$

(1) Sketch the PDFs $f_X(x)$ and $f_Y(y)$ on the same axes.

(2) What is $P[-1 < X \leq 1]$?

(3) What is $P[-1 < Y \leq 1]$?

(4) What is $P[X > 3.5]$?

(5) What is $P[Y > 3.5]$?

Conditioning on a Continuous RV

Definition 3.15 Conditional PDF given an Event

For a random variable X with PDF $f_X(x)$ and an event $B \subset S_X$ with $P[B] > 0$, the conditional PDF of X given B *is*

$$f_{X|B}(x) = \begin{cases} \dfrac{f_X(x)}{P[B]} & x \in B, \\ 0 & \textit{otherwise.} \end{cases}$$

$f_{X|B}(x)\,\Delta x$

$= P[X \le x + \Delta x | B]$

$- P[X \le x | B]$

- Note that the conditional pdf can be written as

$$f_{X|B}(x)\,dx = P[x \le X < x + dx | B]$$

$$= \frac{P[x \le X < x + dx, B]}{P[B]}$$

$X \sim f_X(x), \quad x \in \mathbb{R} \qquad X \in (-\infty, x)$

B

$$f_{X|B}(x) = \frac{d}{dx} P[X \le x | B]$$

$$f_{X|B}(x) \approx \frac{P[X \le x + \Delta x | B] - P[X \le x | B]}{\Delta x}$$

$$f_X(x)\,\Delta x = \frac{P[x \le X < x + \Delta x | B]}{\Delta x}$$

$(x, x+\Delta x)$

$$\frac{P[x \le X \le x + \Delta x, B]}{P[B]\,\Delta x}$$

$x \in B$

Conditioning on a Continuous RV

- What if the event {x <= X < x + dx} is in B?

$$f_{X|B}(x)\,dx = \frac{P[x \le X < x + dx]}{P[B]}$$
$$= \frac{f_X(x)\,dx}{P[B]}$$

- = 0, when event {x <= X < x + dx} is **NOT** in B
- The definition comprises of the above two cases

Theorem 3.23

$$P[X \in (x, x+\Delta x)] = \sum_i P[X \in (x, x+\Delta x), B_i]$$

Given an event space $\{B_i\}$ and the conditional PDFs $f_{X|B_i}(x)$,

$$f_X(x) = \sum_i f_{X|B_i}(x) P[B_i].$$

$v = 0$

a

$a\Delta t$

$$\frac{P[x \le X < x+\Delta x]}{\Delta x}$$

$$\Delta x \, P[x \le X < x+\Delta x \mid B_i]$$

$P[B_i]$

$X \in (x, x+\Delta x)$

$$F(x_0+\Delta x) - F(x_0) \approx \Delta x\, F'(x_0)$$

$$\left.\frac{dF(x)}{dx}\right|_{x=x_0}$$

$$F(x_0+\Delta x) = F(x_0) + \Delta x\, F'(x_0)$$

Example 3.32 Problem

Continuing Example 3.3, when symbol "0" is transmitted (event B_0), X is the Gaussian $(-5, 2)$ random variable. When symbol "1" is transmitted (event B_1), X is the Gaussian $(5, 2)$ random variable. Given that symbols "0" and "1" are equally likely to be sent, what is the PDF of X?

$$f_{X|B_0}(x) = \frac{1}{\sqrt{2\pi(2)^2}}\, e^{-\left[(x-(-5))^2/(2(2)^2)\right]}$$

$$f_{X|B_1}$$

$$f_X(x) = f_{X|B_0}(x)\, P[B_0] + f_{X|B_1}(x)\, P[B_1] \qquad x \in R$$

Conditional Expected Value

Definition 3.16 Given an Event

If $\{x \in B\}$, *the conditional expected value of* X *is*

$$E[X|B] = \int_{-\infty}^{\infty} x f_{X|B}(x)\, dx.$$

$E[X]$

$Y = g(X)$

$E[Y|B]$

$= \int_{-\infty} g(x) f_X(x)\, dx$

Y=g(X), where X is a Continuous Random Variable

- We want to find $f_Y(y)$
- You want a complete probability description of Y
  - Just the expectation or some moment of Y is not enough
- Two step procedure
  - Find the CDF $F_Y(y)$
  - Find $f_Y(y)$. How??

$$Y = a^2X + b + 2b^2$$

$$Y = aX$$

$$P[Y \le y] = P[aX \le y] = P[X \le y/a]$$

$$= F_X(y/a)$$

Scaling a Random Variable

Theorem 3.19

If $Y = aX$, where $a > 0$, then Y has CDF and PDF

$$F_Y(y) = F_X(y/a), \qquad f_Y(y) = \frac{1}{a} f_X(y/a).$$

- We want P[Y<=y]
- P[Y <= y] = P[aX <= y] = P[X <= y/a] = F_X(y/a)

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(y/a)$$

$$= \frac{d}{dx} F_X(y/a) \frac{d}{dy}(y/a) = \frac{1}{a} f_X(y/a)$$

Adding a Constant

Theorem 3.21

If $Y = X + b$,

$$F_Y(y) = F_X(y-b), \qquad f_Y(y) = f_X(y-b).$$

- We want P[Y<=y]
- P[Y <= y] = P[X + b <= y] = P[X <= y-b] = F_X(y-b)
- f_Y(y) = f_X(y-b)

Example 3.24 Problem

Let X be a random variable with CDF $F_X(x)$. Let Y be the output of a clipping circuit with the characteristic $Y = g(X)$ where

$P[Y \le 3] =$

$P[Y \le y]$

$P[Y \le 1]$

$$g(x) = \begin{cases} 1 & x \le 0, \\ 3 & x > 0. \end{cases}$$

$= P[X \le 0] = F_X(0)$

Express $F_Y(y)$ and $f_Y(y)$ in terms of $F_X(x)$ and $f_X(x)$.

$P[Y \le 2] = F_X(0)$ $\quad P[Y < 1] = 0$

Detour...

Definition 3.12 *Unit Impulse (Delta) Function*

Let

$$d_\epsilon(x) = \begin{cases} 1/\epsilon & -\epsilon/2 \le x \le \epsilon/2, \\ 0 & \textit{otherwise}. \end{cases}$$

The unit impulse function *is*

$$\delta(x) = \lim_{\epsilon \to 0} d_\epsilon(x).$$

Delta Function

Figure 3.7

As $\epsilon \to 0$, $d_\epsilon(x)$ approaches the delta function $\delta(x)$. For each ϵ, the area under the curve of $d_\epsilon(x)$ equals 1.

Properties of the Delta Function

- The delta function δ(x) is infinite at x=0 and zero everywhere else.
- The area under the delta function is unity

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1$$

- It has the sifting property. For a continuous function f(x)

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\, dx = f(x_0)$$

- This is because $\delta(x-x_0)$ is δ(x) shifted by x_0. Specifically, we can write

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(x) d_\epsilon(x - x_0)\, dx = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_0 - \epsilon/2}^{x_0 + \epsilon/2} f(x)\, dx$$

Properties of the Delta Function

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(x) d_\epsilon(x - x_0)\, dx = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{x_0 - \epsilon/2}^{x_0 + \epsilon/2} f(x)\, dx$$

- The integral on the right is the average of the function f(x) over the interval of length Є
- The average -> the constant $f(x_0)$ as Є -> 0
- Thus

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0)\, dx = f(x_0)$$

The Unit Step Function

Definition 3.13 *Unit Step Function*

The unit step function *is*

$$u(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0. \end{cases}$$

Note that u(0) = 1. This is by definition and leads to a jump (discontinuity) at x=0

Theorem 3.17

$$\int_{-\infty}^{x} \delta(v)\, dv = u(x).$$

In other words the delta function is the derivative of the unit step function

CDF of a Discrete RV

$$F_X(x) = \sum_{x_i \in S_X} P_X(x_i) u(x - x_i)$$

For any given x, $u(x - x_i) = 1$ for $x >= x_i$

PDF of a Discrete RV

- Having defined the unit step, we can now define a pdf for a discrete random variable!

Delta functions at all x_i are shown of the same height as they all are infinite at the x_i.

Write $P[X = x_i]$ to denote the size of the corresponding step in the CDF.

- Jumps in CDF correspond to delta functions in the pdf!

$Y = X^2$

- X has a discrete uniform distribution and the range $S_X = \{1,2,3,4,5,6\}$
- What is $F_Y(y)$ and $f_Y(y)$?
  - $S_Y = \{1,4,9,16,25,36\}$
  - $P_Y(y) = 1/6$, the pdf will contain impulses with weight 1/6 at each point in S_Y
- X has a discrete uniform distribution and the range $S_X = \{-2, -1, 0, 1,2,3\}$
- What is $F_Y(y)$ and $f_Y(y)$?
  - $S_Y = \{0, 1, 4, 9\}$
  - $P_Y(y) = 1/6, 1/3, 1/3, 1/6$ for Y = 0,1,4, and 9 respectively

$Y = X^2$

- For a continuous RV X
- $P[Y <= y] = P[X^2 <= y] = 0$, for $y < 0$.
- For $y >= 0$ we have

$$P[Y \leq y] = P[X^2 \leq y] = P[-\sqrt{y} \leq X \leq \sqrt{y}]$$
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

- Therefore

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}}, y > 0$$

- $f_Y(y) = 0$ for $y <= 0$.

$y \in (0,1) \rightarrow P[Y \leq y]$

$y \in (1,9)$

$P[Y \in (0,y)]$

$P[Y \le y]$ y $P[Y \le 0]$

$P[Y \le 9]$ $X \in (-\sqrt{9}, \sqrt{9})$

Example 3.26 Problem

Suppose X is uniformly distributed over $[-1, 3]$ and $Y = X^2$. Find the CDF $F_Y(y)$ and the PDF $f_Y(y)$.

- We know

$$P[Y \le y] = P[X^2 \le y] = P[-\sqrt{y} \le X \le \sqrt{y}]$$
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

- Therefore

$$P[Y \le y] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\, dx$$

$Y = X^2$

- However X is Uniform over [-1,3]
  - While [-1,1] is a subset of S_X, the elements in the set (1,3] have no corresponding negative values

- Therefore, for y > 0

$$P[Y \leq y] = \int_{\max(-1,-\sqrt{y})}^{\sqrt{y}} f_X(x)\, dx$$

- For y < 0, we have P[Y <= y] = 0.

$$P[Y \leq y] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\, dx = \frac{\sqrt{y}}{2}, 0 \leq y \leq 1$$

$Y = X^2$

- Also

$$P[Y \leq y] = \int_{-1}^{\sqrt{y}} f_X(x)\, dx = \frac{\sqrt{y}+1}{4}, 1 < y \leq 9$$

Problem 3.6.8

With probability 0.7, the toss of an Olympic shot-putter travels $D = 60 + X$ feet, where X is an exponential random variable with expected value $\mu = 10$. Otherwise, with probability 0.3, a foul is committed by stepping outside of the shot-put circle and we say $D = 0$. What are the CDF and PDF of random variable D?

$P[D \le 59.9]$

$P[D \le d] = 0 \qquad d < 0$

$P[D \le d] = P[D < 0] + P[D = 0] \qquad d \le 0$

$= 0.3$

$T=0 \Leftrightarrow \{L \geq 5\}$

Problem 3.6.9

For 70% of lectures, Professor Y arrives on time. When Professor Y is late, the arrival time delay is a continuous random variable uniformly distributed from 0 to 10 minutes. Yet, as soon as Professor Y is 5 minutes late, all the students get up and leave. (It is unknown if Professor Y still conducts the lecture.) If a lecture starts when Professor Y arrives and always ends 80 minutes after the scheduled starting time, what is the PDF of T, the length of time that the students observe a lecture.

$P[T<0]=0$

$P[T=0]=0.15$

$(0.3)(0.5)$

Problem 3.2.5

For constants a and b, random variable X has PDF

$$f_X(x) = \begin{cases} ax^2 + bx & 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

What conditions on a and b are necessary and sufficient to guarantee that $f_X(x)$ is a valid PDF?

Problem 3.4.14

This problem outlines the steps needed to show that a nonnegative continuous random variable X has expected value

$$E[X] = \int_0^\infty x f_X(x)\, dx = \int_0^\infty [1 - F_X(x)]\, dx.$$

Problem 3.5.10

In mobile radio communications, the radio channel can vary randomly. In particular, in communicating with a fixed transmitter power over a "Rayleigh fading" channel, the receiver signal-to-noise ratio Y is an exponential random variable with expected value γ. Moreover, when $Y = y$, the probability of an error in decoding a transmitted bit is $P_e(y) = Q(\sqrt{2y})$ where $Q(\cdot)$ is the standard normal complementary CDF. The average probability of bit error, also known as the bit error rate or BER, is

$$\overline{P}_e = E\left[P_e(Y)\right] = \int_{-\infty}^{\infty} Q(\sqrt{2y}) f_Y(y)\, dy.$$

Find a simple formula for the BER $\overline{P}_e$ as a function of the average SNR γ.

Birthday Problem (Example 3.5 RN)

- Assume that birthdays occur uniformly over all 365 days of the year. In a group of k people what is the probability that no two people have the same birthday?
- HW: Plot this probability as a function of k
- Probability is the ratio of the number of favorable outcomes to total number of possibilities, which is

$$\frac{(n)_k}{n^k}$$

$$\frac{365 \times 364 \times \cdots (365-k+1)}{(365)^k}$$

Coincidences

- Generalization of the birthday problem
- You have n possible events
  - n=365 in birthday problem
- In a set of k events, the probability that no two events are the same is

$$\frac{(n)_k}{n^k} = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right)$$

- Fix k and let n -> ∞
  - The probability goes to 1!
- Note as n becomes large for a fixed k, it does not matter whether you permute with or without replacement
  - Even when you permute with replacement the probability of picking the same item from a large number n of items is small for small k

Monty Hall Problem

$\rightarrow P[\text{MH picks } 3 \mid \text{Car behind } 2, \text{ You pick } 1]$

$P[\text{Car is behind } 2 \mid \text{You pick } 1]$

http://mathworld.wolfram.com/MontyHallProblem.html

- Assume that a room is equipped with three doors. Behind two are goats, and behind the third is a shiny new car. You are asked to pick a door, and will win whatever is behind it. Let's say you pick door 1.
- Before the door is opened, however, someone who knows what's behind the doors (Monty Hall) opens *one of the other* two doors, revealing a goat, and asks you if you wish to change your selection to the third door (i.e., the door which neither you picked nor he opened). The Monty Hall problem is deciding whether you do.

$P[1] = P[2] = P[3] = 1/3$

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$

You pick

Picks

$$P[\text{Car is behind } 2 \mid \text{You pick } 1, \text{ MH picks } 3]$$

$$= \frac{P[\text{Car is behind } 2, \text{ MH picks } 3 \mid \text{You pick } 1]}{P[\text{MH picks } 3 \mid \text{You pick } 1]}$$

Summary For a RV X

- CDF is given by $F_X(x) = P[X <= x]$
  - Nondecreasing
  - Starts at 0 and ends at 1
  - Is continuous for a continuous RV
  - Contains steps for a discrete RV. Step size at x is the PMF at x, $P_X(x) = P[X=x]$
- PDF $f_X(x)$ is the derivative of $F_X(x)$
  - Always >= 0
  - It is the slope of the CDF at x
  - Area under PDF is 1

$$
\begin{aligned}
P[x_1 \leq X \leq x_2] &= P[X \leq x_2] - P[X < x_1] \\
&= F_X(x_2) - F_X(x_1) \\
&= \int_{x_1}^{x_2} f_X(x)\, dx
\end{aligned}
$$

Summary For a RV X

- Moments of RV X

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

- When conditioning on an event E

$$P_X(x) = P_{X|B}(x)P[B] + P_{X|B^c}(x)P[B^c]$$

$$F_X(x) = F_{X|B}(x)P[B] + F_{X|B^c}(x)P[B^c]$$

$$f_X(x) = f_{X|B}(x)P[B] + f_{X|B^c}(x)P[B^c]$$

$$E[g(x)] = E[g(x)|B]P[B] + E[g(x)|B^c]P[B^c]$$

Pairs of Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

$s \in S$

$X(s)$

$(-\infty,\infty)\times(-\infty,\infty)$

$\{X \le x, Y \le y\}$

We started with…

- An Experiment that contains a
  - Procedure, Observations, and Model
- Later we mapped outcomes to numbers
  - One or more outcomes to a point on the line
- We called the numbers random variables
- We will now map an outcome to a pair of numbers
  - One or more outcomes to a point on a plane
- The numbers in the pair correspond to RVs X and Y
- For example, a transmitted sinusoid that is received with a random amplitude and random phase
  - Outcomes can be described by X = amplitude and Y = phase

Pair of RVs

- We defined a CDF for a single RV X
- For the pair of RVs we define a ***joint*** CDF

Joint Cumulative Distribution Function (CDF)

Definition 4.1

The joint cumulative distribution function *of random variables* X *and* Y *is*

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y].$$

Joint CDF

Figure 4.1

The area of the (X, Y) plane corresponding to the joint cumulative distribution function $F_{X,Y}(x, y)$.

- We are interested in the probability of the intersection of the events {X<=x} and {Y<=y}

Joint CDF

S_Y

Theorem 4.1

For any pair of random variables, X, Y,

$\{X \le x\}$ $\{Y \le y, Y > y\}$

$\{X \le x \cap Y \le y\} \cup$

(a) $0 \le F_{X,Y}(x, y) \le 1$

(b) $F_X(x) = F_{X,Y}(x, \infty)$, $\{X \le x, Y \le \infty\} = \{X \le x\}$ $\{X \le x, Y > y\}$ $P[X \le x, Y < \infty] P[Y \le \infty] = P[X \le x]$

(c) $F_Y(y) =$,

$\{X \le -\infty, Y \le y\}$

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) =$

$P[X \le x, Y \le y]$ $P[X \le x_1, Y \le y_1]$

$A \subseteq B$

$P[A] \le P[B]$

(e) If $x \le x_1$ and $y \le y_1$, then $F_{X,Y}(x, y) \le F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = 1$.

(g) $F_{X,Y}(\infty, -\infty) = ?$

Joint CDF

Theorem 4.1

For any pair of random variables, X, Y,

(a) $0 \leq F_{X,Y}(x, y) \leq 1$,

(b) $F_X(x) = F_{X,Y}(x, \infty)$,

(c) $F_Y(y) = F_{X,Y}(\infty, y)$,

$P[X \leq -\infty, Y \leq y]$

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$,

(e) If $x \leq x_1$ and $y \leq y_1$, then $F_{X,Y}(x, y) \leq F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = 1$.

(g) $F_{X,Y}(\infty, -\infty) = 0$.

Joint PMF

$\frac{a}{b}$

(0,1)

Joint Probability Mass Function

Definition 4.2 *(PMF)*

$S_X \times S_Y$

The joint probability mass function *of discrete random variables* X *and* Y *is*

$$P_{X,Y}(x, y) = P[X = x, Y = y].$$

- This is the probability of the intersection of the events {X=x} and {Y=y}
- We are saying that an observation was made that maps to X=x and Y=y

Example of a Joint PMF

Example 4.1 Problem

$\{a, r\}$

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF of X and Y.

$S_X = \{0, 1, 2\}$ $S_Y = \{0, 1, 2\}$

- What are the possible observations (sequence of observations)?
  - {aa, ar, ra, rr}

$S_X \times S_Y$

Example of a Joint PMF

Example 4.1 Problem

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF of X and Y.

- aa => {X = 2, Y = 2} ; P[X = 2, Y = 2] = 0.81
- ar => {X = 1, Y = 1} ; P[X = 1, Y = 1] = 0.09
- ra => {X = 1, Y = 0} ; P[X = 1, Y = 0] = 0.09
- rr => {X = 0, Y = 0} ; P[X = 0, Y = 0] = 0.01

Example of a Joint PMF

Three possible representations of a Joint PMF

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	0.01	0	0
$x = 1$	0.09	0.09	0
$x = 2$	0	0	0.81

$$P_{X,Y}(x, y) = \begin{cases} 0.81 & x = 2, y = 2, \\ 0.09 & x = 1, y = 1, \\ 0.09 & x = 1, y = 0, \\ 0.01 & x = 0, y = 0. \\ 0 & \text{otherwise} \end{cases}$$

Joint PMF

$$\sum_{x \in S_X} \sum_{y \in S_Y} P_{X,Y}(x,y) = ?$$

- 1. Because you are summing over the probabilities of all mutually exclusive events {x,y} in S
- In other words, you are saying that at least one pair of {x,y} will be observed where x is in range of X and y in range of Y
- Note that the event {X=x1, Y=y1} is mutually exclusive of the event {X=x2, Y=y2} if x1 ≠ x2 or y1 ≠ y2

Probability of an Event B that is in the set $S_X \times S_Y$

Theorem 4.2

For discrete random variables X and Y and any set B in the X, Y plane, the probability of the event $\{(X, Y) \in B\}$ is

$$P[B] = \sum_{(x,y)\in B} P_{X,Y}(x, y).$$

VVVS Problem

Quiz 4.2

The joint PMF $P_{Q,G}(q,g)$ for random variables Q and G is given in the following table:

$\{Q=0\} = (\{Q=0\} \cap \{G=0\}) \cup$

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

(4.12)

Calculate the following probabilities:

(1) $P[Q=0]$

(2) $P[Q=G]$

(3) $P[G>1]$

(4) $P[G>Q]$

The event B in Theorem 4.2 can be defined to be any of the events in (1) to (4).

Marginal PMF

- For discrete RVs X and Y with joint PMF $P_{X,Y}(x,y)$, $P_X(x)$ and $P_Y(y)$ are defined as the marginal PMFs of X and Y respectively

$P[X=x] = \sum_{y \in S_Y} P[X=x, Y=y]$

$P[X=x,Y=y]$

$P[Y=y]$

$P[X=x] = \sum_{y \in S_Y} P[X=x, Y=y]$

Quiz 4.3

The probability mass function $P_{H,B}(h,b)$ for the two random variables H and B is given in the following table. Find the marginal PMFs $P_H(h)$ and $P_B(b)$.

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$
$h=-1$	0	0.4	0.2
$h=0$	0.1	0	0.1
$h=1$	0.1	0.1	0

$P[H=-1]$ —

$P[H=0]$ — (4.20)

$P[H=1]$ —

$P[\text{Any other } H] = 0$

Marginal PMF

Theorem 4.3

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$,

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y), \qquad P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y).$$

- We looked at the joint CDF and the joint PMF. What do you expect next?

$$\underbrace{\frac{P[x \le X < x+\Delta x,\ y \le Y < y+\Delta y]}{(\Delta x \Delta y)}}_{F_{X,Y}(x,y)} \underset{\Delta x, \Delta y \to 0}{=} \frac{P[X \le x+\Delta x,\ Y \le y+\Delta x] - P[X \le x,\ Y \le y+\Delta y] - P[X \le x+\Delta x,\ Y \le y] + P[X \le x,\ Y \le y]}{(\Delta x \Delta y)}$$

$F_{X,Y}(x+\Delta x, y+\Delta y)$

Joint PDF

Joint Probability Density Function

Definition 4.3

The joint PDF of the continuous random variables X and Y is a function $f_{X,Y}(x, y)$ with the property

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du.$$

Theorem 4.4

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

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy$$

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

(a) $f_{X,Y}(x, y) \geq 0$ for all (x, y),

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1.$

- Nonnegative
- Area under it is unity

Theorem 4.5

Theorem 4.5

$$P\left[x_1 < X \leq x_2, y_1 < Y \leq y_2\right] = F_{X,Y}\left(x_2, y_2\right) - F_{X,Y}\left(x_2, y_1\right) - F_{X,Y}\left(x_1, y_2\right) + F_{X,Y}\left(x_1, y_1\right)$$

- Note that the theorem is valid since we assumed that X takes values along the x-axis and Y along the y-axis.
  - Outcome {X=x,Y=y} is a point in the *xy-plane*

Properties of a Joint PDF

- For a point (x,y) in the xy plane, we know that

$$\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)}{\Delta x \Delta y}$$

- Therefore we can write

$$f_{X,Y}(x,y)\Delta x \Delta y \approx F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)$$

- Hence

$$f_{X,Y}(x,y)\Delta x \Delta y \approx P[x < X \leq x+\Delta x, y < Y \leq y+\Delta y]$$

Using a Joint PDF

- Even when the event corresponds to a rectangular area calculating probability of the event using the CDF is nontrivial!
- Using a PDF is usually more convenient

Theorem 4.7

$z = x^2 + y^2$

The probability that the continuous random variables (X, Y) are in A is

$$P[A] = \iint_A f_{X,Y}(x, y)\, dx\, dy.$$

R_2

- A could be the rectangular region we looked at earlier
- A can be an area of any shape or size

VVVS Problem

Example 4.4 Problem

Random variables X and Y have joint PDF

$$f_{X,Y}(x, y) = \begin{cases} c & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}$$

Find the constant c and $P[A] = P[2 \le X < 3, 1 \le Y < 3]$.

- c=1/15
- P[A] = 2/15
- Note that A is a rectangular area and so the problem is very straightforward

Calculating a CDF from a PDF!

Example 4.5 Problem

$P(X \le x, Y \le y) = 1 \quad x > 1, y > 1$

$x > 1, 0 \le y < 1$

Find the joint CDF $F_{X,Y}(x, y)$ when X and Y have joint PDF

$P[X \le x, Y \le y] \quad \int_0^1 \int_0^y dy\, dx$

$$f_{X,Y}(x, y) = \begin{cases} 2 & 0 \le y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

$\iint_A f_{X,Y}(x,y)\, dx\, dy \qquad 2(y)(1-y) + 2\left(\frac{1}{2}y^2\right) = 2y - 2y^2 + y^2$

$2 \iint_A dx\, dy$

- All valid outcomes lie in the triangle

VS Problem

Quiz 4.4

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x, y) = \begin{cases} cxy & 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.33}$$

Find the constant c. What is the probability of the event $A = X^2 + Y^2 \le 1$?

- We want to find the probability that a pair (x,y) lies inside the unit circle
- Use polar coordinates
  - dx dy = r dr dϕ

Marginal PDF

Theorem 4.8

If X and Y are random variables with joint PDF $f_{X,Y}(x, y)$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx.$$

- Note that P[X <= x] = P[X<=x, Y<∞]
- We have

$$P[X \leq x] = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X,Y}(u, v)\, dv\, du$$

- Taking the derivative with respect to x (and using the fundamental theorem of calculus) we get our desired result

VVVS Problem

Quiz 4.5

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x, y) = \begin{cases} 6(x + y^2)/5 & 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.40}$$

Find $f_X(x)$ and $f_Y(y)$, the marginal PDFs of X and Y.

- You must get

$$f_X(x) = \begin{cases} \frac{6x+2}{5} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3+6y^2}{5} & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Functions of Two Random Variables

- Assume life was easier and you had just a midterm (30% weight) and a final exam.
- Let X be the RV that denotes your midterm marks
  - $S_X = [0,30]$
- Let Y be the RV that denotes your final exam marks
  - $S_Y = [0,70]$
- Let G be the RV that denotes your grade
- $G = g(X,Y)$
  - Where g is determined by yours truly
  - Example $G = 0.3\,X + 0.7\,Y$

A More Useful Example

- More than one antenna at your receiver receives a transmitted signal

Pic courtesy http://www.dsplog.com/2008/09/06/receiver-diversity-selection-diversity/

- For N=2, we have received signals (RVs) R_1 and R_2
- The resultant received signal $R = w_1R_1 + w_2R_2$
- We could keep it simpler and only select the stronger signal
  - $R = \max(R_1,R_2)$

Functions of Two Random Variables

- We want to find the probability model that corresponds to the resulting RV
- **We know the joint probability model**
- W = X + Y
  - PMF, CDF, PDF of W?
- As always if X and Y are discrete, life is easy

Theorem 4.9

For discrete random variables X and Y, the derived random variable $W = g(X, Y)$ has PMF

$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x, y).$$

Continuous Function of Continuous RVs

- W = g(X,Y)
  - g is continuous and so are X and Y.
- W is thus a continuous RV

Theorem 4.10

For continuous random variables X and Y, the CDF of $W = g(X, Y)$ is

$$F_W(w) = P[W \leq w] = \iint_{g(x,y) \leq w} f_{X,Y}(x, y)\, dx\, dy.$$

X, Y

$\{X \le w, Y \le w\}$

$W = \max(X, Y)$

$\{W \le w\}$

$P[\{W > w\}]$ $F_W(w)$

$= P[X > w, Y > w]$

$F_W(w) = F_{X,Y}(w, w)$

$P[W \le w] = 1 - P[X > w, Y > w]$

Maximum of Two RVs

Theorem 4.11

For continuous random variables X and Y, the CDF of $W = \max(X, Y)$ is

$$F_W(w) = F_{X,Y}(w, w) = \int_{-\infty}^{w} \int_{-\infty}^{w} f_{X,Y}(x, y)\, dx\, dy.$$

What if W = min(X,Y)?

$$P[W > w] = P[X > w, Y > w]$$

Expected Values!

$$E[W] = \sum_{w \in S_W} w \, P[W=w]$$

Theorem 4.12

For random variables X and Y, the expected value of $W = g(X, Y)$ is

Discrete: $$E[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y}(x, y)$$

$$\sum_{w \in S_W} w \sum_{g(x,y)=w} \sum P[X=x, Y=y]$$

Continuous: $$E[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) \, dx \, dy.$$

Theorem 4.13

$$E[g_1(X, Y) + \cdots + g_n(X, Y)] = E[g_1(X, Y)] + \cdots + E[g_n(X, Y)].$$

- A very useful theorem!

Suppose $g(X,Y) = X = W$

$$E[W] = E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \, f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} x \underbrace{\left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right]}_{f_X(x)} dx$$

$$\boxed{E[W] = \sum_{w \in S_W} w \, P[W=w]}$$

$$= \sum_{w \in S_W} w \boxed{\sum_{\{g(x,y)=w\}} \sum P[X=x, Y=y]}$$

$$\{(x,y) : g(x,y) = w\}$$

$$E[W] = \int_{-\infty}^{\infty} w \, f_W(w) \, dw$$

$$\iint_{g(x,y)=w} f_{X,Y}(x,y) \, dx \, dy$$

Expected Values

- E[X + Y] = ?
- $E[(X + Y)^2]$ = ?

Theorem 4.15

The variance of the sum of two random variables is

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2E[(X-\mu_X)(Y-\mu_Y)].$$

Definition 4.4 Covariance

The covariance *of two random variables* X *and* Y *is*

$$\mathrm{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)].$$

Pairs of Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

$s \in S$

$X(s)$

$(-\infty,\infty)\times(-\infty,\infty)$

$\{X \le x, Y \le y\}$

We started with…

- An Experiment that contains a
  - Procedure, Observations, and Model
- Later we mapped outcomes to numbers
  - One or more outcomes to a point on the line
- We called the numbers random variables
- We will now map an outcome to a pair of numbers
  - One or more outcomes to a point on a plane
- The numbers in the pair correspond to RVs X and Y
- For example, a transmitted sinusoid that is received with a random amplitude and random phase
  - Outcomes can be described by X = amplitude and Y = phase

Pair of RVs

- We defined a CDF for a single RV X
- For the pair of RVs we define a ***joint*** CDF

Joint Cumulative Distribution Function (CDF)

Definition 4.1

The joint cumulative distribution function *of random variables* X *and* Y *is*

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y].$$

Joint CDF

Figure 4.1

The area of the (X, Y) plane corresponding to the joint cumulative distribution function $F_{X,Y}(x, y)$.

- We are interested in the probability of the intersection of the events {X<=x} and {Y<=y}

Joint CDF

Theorem 4.1

For any pair of random variables, X, Y,

$\{X \le x\}$ $\{Y \le y, Y > y\}$

$\{X \le x \cap Y \le y\} \cup$

(a) $\boxed{0} \le F_{X,Y}(x, y) \le \boxed{1}$

(b) $F_X(x) = \boxed{F_{X,Y}(x, \infty)}$, $\{X \le x, Y \le \infty\} = \{X \le x\}$ $\{X \le x, Y > y\}$ $P[X \le x, Y < \infty] P[Y \le \infty] = P[X \le x]$

(c) $F_Y(y) = \boxed{}$,

$\{X \le -\infty, Y \le y\}$

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = \boxed{}$

$P[X \le x, Y \le y]$ $P[X \le x_1, Y \le y_1]$

$A \subseteq B$

$P[A] \le P[B]$

(e) If $x \le x_1$ and $y \le y_1$, then $F_{X,Y}(x, y) \boxed{\le} F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = \boxed{1}$.

(g) $F_{X,Y}(\infty, -\infty) = ?$

Joint CDF

Theorem 4.1

For any pair of random variables, X, Y,

(a) $0 \leq F_{X,Y}(x, y) \leq 1$,

(b) $F_X(x) = F_{X,Y}(x, \infty)$,

(c) $F_Y(y) = F_{X,Y}(\infty, y)$,

$P[X \leq -\infty, Y \leq y]$

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$,

(e) If $x \leq x_1$ and $y \leq y_1$, then $F_{X,Y}(x, y) \leq F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = 1$.

(g) $F_{X,Y}(\infty, -\infty) = 0$.

Joint PMF

$\frac{a}{b}$

$(0,1)$

Joint Probability Mass Function

$S_X \times S_Y$

Definition 4.2 (PMF)

The joint probability mass function *of discrete random variables* X *and* Y *is*

$$P_{X,Y}(x, y) = P[X = x, Y = y].$$

- This is the probability of the intersection of the events {X=x} and {Y=y}
- We are saying that an observation was made that maps to X=x and Y=y

Example of a Joint PMF

Example 4.1 Problem

$\{a, r\}$

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF of X and Y.

$S_X = \{0, 1, 2\}$ $S_Y = \{0, 1, 2\}$

- What are the possible observations (sequence of observations)?
  - {aa, ar, ra, rr}

$S_X \times S_Y$

Example of a Joint PMF

Example 4.1 Problem

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF of X and Y.

- aa => {X = 2, Y = 2} ; P[X = 2, Y = 2] = 0.81
- ar => {X = 1, Y = 1} ; P[X = 1, Y = 1] = 0.09
- ra => {X = 1, Y = 0} ; P[X = 1, Y = 0] = 0.09
- rr => {X = 0, Y = 0} ; P[X = 0, Y = 0] = 0.01

Example of a Joint PMF

Three possible representations of a Joint PMF

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	0.01	0	0
$x = 1$	0.09	0.09	0
$x = 2$	0	0	0.81

$$P_{X,Y}(x, y) = \begin{cases} 0.81 & x = 2, y = 2, \\ 0.09 & x = 1, y = 1, \\ 0.09 & x = 1, y = 0, \\ 0.01 & x = 0, y = 0. \\ 0 & \text{otherwise} \end{cases}$$

Joint PMF

$$\sum_{x \in S_X} \sum_{y \in S_Y} P_{X,Y}(x,y) = ?$$

- 1. Because you are summing over the probabilities of all mutually exclusive events {x,y} in S
- In other words, you are saying that at least one pair of {x,y} will be observed where x is in range of X and y in range of Y
- Note that the event {X=x1, Y=y1} is mutually exclusive of the event {X=x2, Y=y2} if x1 ≠ x2 or y1 ≠ y2

Probability of an Event B that is in the set S_X ***x*** S_Y

Theorem 4.2

For discrete random variables X and Y and any set B in the X, Y plane, the probability of the event $\{(X, Y) \in B\}$ is

$$P[B] = \sum_{(x,y)\in B} P_{X,Y}(x, y).$$

VVVS Problem

Quiz 4.2

The joint PMF $P_{Q,G}(q,g)$ for random variables Q and G is given in the following table:

$\{Q=0\} = (\{Q=0\} \cap \{G=0\}) \cup$

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

(4.12)

Calculate the following probabilities:

(1) $P[Q=0]$

(2) $P[Q=G]$

(3) $P[G>1]$

(4) $P[G>Q]$

The event B in Theorem 4.2 can be defined to be any of the events in (1) to (4).

Marginal PMF

- For discrete RVs X and Y with joint PMF $P_{X,Y}(x,y)$, $P_X(x)$ and $P_Y(y)$ are defined as the marginal PMFs of X and Y respectively

$P[X=x, Y=y]$

$P[X=x] = \sum_{y \in S_Y} P[X=x, Y=y]$

$P[Y=y]P[X=x] = \sum_{y \in S_Y} P[X=x, Y=y]$

Quiz 4.3

The probability mass function $P_{H,B}(h,b)$ for the two random variables H and B is given in the following table. Find the marginal PMFs $P_H(h)$ and $P_B(b)$.

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$
$h=-1$	0	0.4	0.2
$h=0$	0.1	0	0.1
$h=1$	0.1	0.1	0

(4.20)

$P[H=-1]$ —

$P[H=0]$ —

$P[H=1]$ —

$P[\text{Any other } H] = 0$

Marginal PMF

Theorem 4.3

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$,

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y), \qquad P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y).$$

- We looked at the joint CDF and the joint PMF. What do you expect next?

$$\frac{P[x \le X < x + \Delta x,\ y \le Y < y + \Delta y]}{(\Delta x\, \Delta y)} \underset{\Delta x, \Delta y \to 0}{\longrightarrow} F_{X,Y}(x, y)$$

$$\frac{P[X \le x + \Delta x,\ Y \le y + \Delta x] - P[X \le x,\ Y \le y + \Delta y] - P[X \le x + \Delta x,\ Y \le y] + P[X \le x,\ Y \le y]}{(\Delta x\, \Delta y)} \underset{\Delta x, \Delta y \to 0}{\longrightarrow} F_{X,Y}(x + \Delta x, y + \Delta y)$$

Joint PDF

Joint Probability Density Function (PDF)

$$\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} F_{X,Y}(x,y)\right) = \lim_{\Delta x \to 0} \frac{F_{X,Y}(x+\Delta x, y) - F_{X,Y}(x,y)}{\Delta x}$$

$$f_{X,Y}(x,y)$$

$$f(x) \qquad f(x_0) \qquad x_0 + \Delta x$$

Definition 4.3

$$f(x_0 + \Delta x) \approx f(x_0) + \Delta x f'(x_0)$$

$$f(x_0+\Delta x, y_0+\Delta y)$$

The joint PDF of the continuous random variables X and Y is a function $f_{X,Y}(x, y)$ with the property

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du.$$

Theorem 4.4

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

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy$$

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

(a) $f_{X,Y}(x, y) \geq 0$ for all (x, y),

(b) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1.$

- Nonnegative
- Area under it is unity

Theorem 4.5

Theorem 4.5

$$P\left[x_1 < X \leq x_2, y_1 < Y \leq y_2\right] = F_{X,Y}\left(x_2, y_2\right) - F_{X,Y}\left(x_2, y_1\right) - F_{X,Y}\left(x_1, y_2\right) + F_{X,Y}\left(x_1, y_1\right)$$

- Note that the theorem is valid since we assumed that X takes values along the x-axis and Y along the y-axis.
  - Outcome {X=x,Y=y} is a point in the *xy-plane*

Properties of a Joint PDF

- For a point (x,y) in the xy plane, we know that

$$\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)}{\Delta x \Delta y}$$

- Therefore we can write

$$f_{X,Y}(x,y)\Delta x \Delta y \approx F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)$$

- Hence

$$f_{X,Y}(x,y)\Delta x \Delta y \approx P[x < X \leq x+\Delta x, y < Y \leq y+\Delta y]$$

Using a Joint PDF

- Even when the event corresponds to a rectangular area calculating probability of the event using the CDF is nontrivial!
- Using a PDF is usually more convenient

Theorem 4.7

$z = x^2 + y^2$

The probability that the continuous random variables (X, Y) are in A is

$$P[A] = \iint_A f_{X,Y}(x, y)\, dx\, dy.$$

- A could be the rectangular region we looked at earlier
- A can be an area of any shape or size

VVVS Problem

Example 4.4 Problem

Random variables X and Y have joint PDF

$$f_{X,Y}(x, y) = \begin{cases} c & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}$$

Find the constant c and $P[A] = P[2 \le X < 3, 1 \le Y < 3]$.

- c=1/15
- P[A] = 2/15
- Note that A is a rectangular area and so the problem is very straightforward

Calculating a CDF from a PDF!

Example 4.5 Problem

$P[X \le x, Y \le y] = 1 \quad x > 1, y > 1$

$x > 1, 0 \le y < 1$

Find the joint CDF $F_{X,Y}(x, y)$ when X and Y have joint PDF

$P[X \le x, Y \le y] \quad \int_0^1 \int_0^y dy\, dx$

$$f_{X,Y}(x, y) = \begin{cases} 2 & 0 \le y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

$\iint_A f_{X,Y}(x,y)\, dx\, dy \qquad 2(y)(1-y) + 2\left(\frac{1}{2}y^2\right) = 2y - 2y^2 + y^2$

$2 \iint_A dx\, dy$

- All valid outcomes lie in the triangle

VS Problem

Quiz 4.4

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x, y) = \begin{cases} cxy & 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.33}$$

Find the constant c. What is the probability of the event $A = X^2 + Y^2 \le 1$?

- We want to find the probability that a pair (x,y) lies inside the unit circle
- Use polar coordinates
  - dx dy = r dr dϕ

Marginal PDF

Theorem 4.8

If X and Y are random variables with joint PDF $f_{X,Y}(x, y)$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx.$$

- Note that P[X <= x] = P[X<=x, Y<∞]
- We have

$$P[X \leq x] = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X,Y}(u, v)\, dv\, du$$

- Taking the derivative with respect to x (and using the fundamental theorem of calculus) we get our desired result

VVVS Problem

Quiz 4.5

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} 6(x+y^2)/5 & 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.40}$$

Find $f_X(x)$ and $f_Y(y)$, the marginal PDFs of X and Y.

- You must get

$$f_X(x) = \begin{cases} \frac{6x+2}{5} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3+6y^2}{5} & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Functions of Two Random Variables

- Assume life was easier and you had just a midterm (30% weight) and a final exam.
- Let X be the RV that denotes your midterm marks
  - $S_X = [0,30]$
- Let Y be the RV that denotes your final exam marks
  - $S_Y = [0,70]$
- Let G be the RV that denotes your grade
- $G = g(X,Y)$
  - Where g is determined by yours truly
  - Example $G = 0.3\,X + 0.7\,Y$

A More Useful Example

- More than one antenna at your receiver receives a transmitted signal

Pic courtesy http://www.dsplog.com/2008/09/06/receiver-diversity-selection-diversity/

- For N=2, we have received signals (RVs) R_1 and R_2
- The resultant received signal $R = w_1R_1 + w_2R_2$
- We could keep it simpler and only select the stronger signal
  - $R = \max(R_1,R_2)$

Functions of Two Random Variables

- We want to find the probability model that corresponds to the resulting RV
- **We know the joint probability model**
- W = X + Y
  - PMF, CDF, PDF of W?
- As always if X and Y are discrete, life is easy

Theorem 4.9

For discrete random variables X and Y, the derived random variable $W = g(X, Y)$ has PMF

$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x, y).$$

Continuous Function of Continuous RVs

- W = g(X,Y)
  - g is continuous and so are X and Y.
- W is thus a continuous RV

Theorem 4.10

For continuous random variables X and Y, the CDF of $W = g(X, Y)$ is

$$F_W(w) = P[W \leq w] = \iint_{g(x,y) \leq w} f_{X,Y}(x, y)\, dx\, dy.$$

X, Y

$\{X \le w, Y \le w\}$

$W = \max(X, Y)$

$\{W \le w\}$

$P[W > w]$ $\quad F_W(w)$

$= P[X > w, Y > w]$

$F_W(w) = F_{X,Y}(w, w)$

$P[W \le w] = 1 - P[X > w, Y > w]$

Maximum of Two RVs

Theorem 4.11

For continuous random variables X and Y, the CDF of $W = \max(X, Y)$ is

$$F_W(w) = F_{X,Y}(w, w) = \int_{-\infty}^{w} \int_{-\infty}^{w} f_{X,Y}(x, y)\, dx\, dy.$$

What if W = min(X,Y)?

$$P[W > w] = P[X > w, Y > w]$$

Expected Values!

$$E[W] = \sum_{w \in S_W} w \, P[W=w]$$

Theorem 4.12

For random variables X and Y, the expected value of $W = g(X, Y)$ is

Discrete: $$E[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y}(x, y)$$

$$\sum_{w \in S_W} w \sum_{g(x,y)=w} \sum P[X=x, Y=y]$$

Continuous: $$E[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) \, dx \, dy.$$

Theorem 4.13

$$E[g_1(X, Y) + \cdots + g_n(X, Y)] = E[g_1(X, Y)] + \cdots + E[g_n(X, Y)].$$

- A very useful theorem!

Suppose $g(X,Y) = X = W$

$$E[W] = E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \, f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} x \underbrace{\left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right]}_{f_X(x)} dx$$

$$\boxed{E[W] = \sum_{w \in S_W} w \, P[W=w]}$$

$$= \sum_{w \in S_W} w \boxed{\sum_{\{g(x,y)=w\}} P[X=x, Y=y]}$$

$$E[W] = \int_{-\infty}^{\infty} w \, f_W(w) \, dw \qquad \{(x,y) : g(x,y) = w\}$$

$$\iint_{g(x,y)=w} f_{X,Y}(x,y) \, dx \, dy$$

Expected Values

- E[X + Y] = ?
- $E[(X + Y)^2]$ = ?

Theorem 4.15

The variance of the sum of two random variables is

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2E[(X-\mu_X)(Y-\mu_Y)].$$

Definition 4.4 Covariance

The covariance *of two random variables* X *and* Y *is*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)\, dx\, dy$$

$$\mathrm{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)].$$

Definition 4.5 Correlation

The correlation *of* X *and* Y *is* $r_{X,Y} = E[XY]$

Theorem 4.16

(a) $\mathrm{Cov}[X, Y] = r_{X,Y} - \mu_X \mu_Y$.

(b) $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X, Y]$.

(c) If $X = Y$, $\mathrm{Cov}[X, Y] = \mathrm{Var}[X] = \mathrm{Var}[Y]$ and $r_{X,Y} = E[X^2] = E[Y^2]$.

Example 4.12 Problem

For the integrated circuits tests in Example 4.1, we found in Example 4.3 that the probability model for X and Y is given by the following matrix.

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$	$P_X(x)$
$x = 0$	0.01	0	0	
$x = 1$	0.09	0.09	0	
$x = 2$	0	0	0.81	
$P_Y(y)$				

(4.73)

Find $r_{X,Y}$ and $\mathrm{Cov}[X, Y]$.

VS Problem

Example 4.12 Problem

For the integrated circuits tests in Example 4.1, we found in Example 4.3 that the probability model for X and Y is given by the following matrix.

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$	$P_X(x)$
$x = 0$	0.01	0	0	0.01
$x = 1$	0.09	0.09	0	0.18
$x = 2$	0	0	0.81	0.81
$P_Y(y)$	0.10	0.09	0.81	

(4.73)

Find $r_{X,Y}$ and $\mathrm{Cov}[X, Y]$.

- 3.33 and 0.252

More Definitions!

Definition 4.6 Orthogonal Random Variables

Random variables X *and* Y *are* orthogonal *if* $r_{X,Y} = 0$.

(x_1, y_1) (x_2, y_2) (x_3, y_3) . . . . (x_n, y_n)

Definition 4.7 Uncorrelated Random Variables

$\vec{x}^T \vec{y} = (x_1 \; x_2 \cdots x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\dfrac{x_1y_1 + x_2y_2 + x_3y_3 + \cdots + x_ny_n}{n}$

Random variables X *and* Y *are* uncorrelated *if* $\mathrm{Cov}[X, Y] = 0$.

$E(XY) = E[X]E[Y]$

Definition 4.8 Correlation Coefficient

The correlation coefficient *of two random variables* X *and* Y *is*

$$\rho_{X,Y} = \frac{\mathrm{Cov}\,[X, Y]}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}\,[X, Y]}{\sigma_X \sigma_Y}.$$

Correlation Coefficient

Theorem 4.17

$$-1 \leq \rho_{X,Y} \leq 1.$$

- Let W = X – aY, for any constant a
- We have Var[W] = Var[X] – 2a Cov[X,Y] + a^2 Var[Y]
- Use the fact that Var[W] >= 0
- Let a = σ_x/σ_y to get the upper bound
- Let a = - σ_x/σ_y to get the lower bound

Correlation Coefficient

- A positive correlation coefficient implies that when X is high wrt E[X], even Y tends to be high wrt E[Y]
  - When we observe high values of X it is likely that Y too is high
- A negative coefficient implies that when X is low, Y is likely to be high and vice-versa
- If X and Y are uncorrelated then no such trend is observed

ρ_{XY}^2

Examples (Scatter Plots)

Weisstein, Eric W. "Correlation Coefficient." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/CorrelationCoefficient.html

Y is an Affine Function of X

Theorem 4.18

If X and Y are random variables such that $Y = aX + b$,

$$\rho_{X,Y} = \begin{cases} -1 & a < 0, \\ 0 & a = 0, \\ 1 & a > 0. \end{cases}$$

$\mu_Y =$
$E[Y]$
$= aE[X] + b$

$E[Y - \mu_Y]$
$= E[Y]$

$E[X(aX+b)]$
$= E[aX^2 + bX]$
$= aE[X^2] + bE[X]$

$$\frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

$$\frac{E[XY] - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

$$\frac{aE[X^2] + bE[X] - E[X](aE[X] + b)}{}$$

Conditioning on an Event

Definition 4.9 *Conditional Joint PMF*

For discrete random variables X and Y and an event, B with $P[B] > 0$, the conditional joint PMF *of X and Y given B is*

$$P_{X,Y|B}(x, y) = P[X = x, Y = y|B].$$

$\{X=x, Y=y\}$

$$P[X=x, Y=y|B] = \frac{P[X=x, Y=y, B]}{P[B]}$$

Theorem 4.19

For any event B, a region of the X, Y plane with $P[B] > 0$,

$$P_{X,Y|B}(x, y) = \begin{cases} \dfrac{P_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

VS Problem

Example 4.13 Problem

$P[X=1, Y=1 | B] = \dfrac{1/4}{\left(\frac{1}{4}+\frac{1}{8}+\frac{1}{8}+\frac{1}{12}\right)}$

Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$ as shown. Let B denote the event $X + Y \leq 4$. Find the conditional PMF of X and Y given B.

$P[B] =$

- P[B] = 7/12

Conditioning on an Event

Definition 4.10 Conditional Joint PDF

Given an event B with $P[B] > 0$, the conditional joint probability density function of X and Y is

$$f_{X,Y|B}(x, y) = \begin{cases} \dfrac{f_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

$$\frac{P[x \le X < x+\Delta x,\ y \le Y < y+\Delta x]}{\Delta x \Delta y}$$

$$\approx \frac{P[x \le X < x+\Delta x,\ y \le Y < y+\Delta y,\ B]}{\Delta x \Delta y}$$

Example 4.14 Problem

X and Y are random variables with joint PDF

$$f_{X,Y}(x, y) = \begin{cases} 1/15 & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.83}$$

Find the conditional PDF of X and Y given the event $B = \{X + Y \ge 4\}$.

Conditioning on an Event

Theorem 4.20 Conditional Expected Value

For random variables X and Y and an event B of nonzero probability, the conditional expected value of $W = g(X, Y)$ given B is

Discrete: $$E[W|B] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y|B}(x, y)$$

Continuous: $$E[W|B] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y|B}(x, y)\, dx\, dy.$$

- Nothing special about this
- You use the conditional PMF and PDF respectively

Conditioning on a RV

Definition 4.12 Conditional PMF

$P[Y=y]$

For any event $Y = y$ *such that* $P_Y(y) > 0$*, the* conditional PMF *of* X *given* $Y = y$ *is*

$$P_{X|Y}(x|y) = P[X = x|Y = y].$$

Theorem 4.22

For random variables X and Y with joint PMF $P_{X,Y}(x, y)$, and x and y such that $P_X(x) > 0$ and $P_Y(y) > 0$,

$$P_{X,Y}(x, y) = P_{X|Y}(x|y) P_Y(y) = P_{Y|X}(y|x) P_X(x).$$

- Proof anyone??

$P[X=x|Y=y]$

VS Problem

Example 4.17 Problem

Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$, as given in Example 4.13 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

$$P[X=1 \mid Y=2] = \frac{P[X=1, Y=2]}{P[Y=2]}$$

- You have the joint PMF
- From it you can calculate the PMF of X and hence the conditional PMF of Y
- How many conditional PMF(s) of Y do we have?

VVVVS Problem

$S_X = \{1, 2, 3, 4\}$

- For a given y

$$P_{Y|X}(y|1) + P_{Y|X}(y|2) + P_{Y|X}(y|3) + P_{Y|X}(y|4) = ?$$

$$P[Y=1 | S_X] + P[Y=2 | S_X] + \cdots + P[Y=4 | S_X]$$

- For a given x

$$P_{Y|X}(1|x) + P_{Y|X}(2|x) + P_{Y|X}(3|x) + P_{Y|X}(4|x) = ?$$

$$P[Y=1 | X=x] + P[Y=2 | X=x] + \cdots + P[Y=4 | X=x]$$

$$\sum_{x \in S_X} P[X=x | Y=y] \qquad \sum_{y \in S_Y} P[X=x | Y=y]\, P[Y=y]$$

$$= \sum_{x \in S_X} \frac{P[X=x, Y=y]}{P[Y=y]} \longrightarrow \sum_{x \in S_X} P[X=x | Y=y]\, P[X=x]$$

Conditional Expectation

$E[P_X(X) | Y=y]$

Theorem 4.23 Conditional Expected Value of a Function

X and Y are discrete random variables. For any $y \in S_Y$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

$$E[g(X, Y)|Y = y] = \sum_{x \in S_X} g(x, y) P_{X|Y}(x|y).$$

$E[E[g(X,Y)|Y]]$

$P_{X|Y}(x|y)$

- The conditional expectation is a function of which variable(s)?
- E[g(X,Y)] is a function of which variable(s)?

$$f_X(x) = \int_{y \in S_Y} f_{X,Y}(x,y)\, dy \longrightarrow \int_{x \in S_X} f_{X|Y}(x|y)\, dx$$

$$\longrightarrow \int_{y \in S_Y} f_{X|Y}(x|y)\, dy$$

Now For the Continuous Case

Definition 4.13 Conditional PDF

For y such that $f_Y(y) > 0$, the conditional PDF of X given $\{Y = y\}$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

Theorem 4.24

$$f_{X,Y}(x, y) = f_{Y|X}(y|x)\, f_X(x) = f_{X|Y}(x|y)\, f_Y(y).$$

Problem

- You are given the conditional pdf $f_{Y|X}(y|x)$ and $f_X(x)$

$$P[Y \le y] = \int_{-\infty}^{y}\int_{-\infty}^{\infty} \underbrace{f_{X,Y}(x,y)}_{f_{Y|X}(y|x)\,f_X(x)}\,dx\,dy$$

- How would you find $F_Y(y)$?

Conditional Expected Value of a Function

Definition 4.14

For continuous random variables X *and* Y *and any* y *such that* $f_Y(y) > 0$, *the* conditional expected value *of* $g(X, Y)$ *given* $Y = y$ *is*

$$E\left[g(X,Y)|Y=y\right] = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)\, dx.$$

$$E[g(X,y)|X=y] \quad E[g(X,y)] = \int g(x,y) f_X(x)\,dx$$

$$\int g(x,y) f_{X|Y}(x|y)\,dx$$

Conditional Expected Value

Definition 4.15 *Conditional Expected Value*

The conditional expected value $E[X|Y]$ is a function of random variable Y such that if $Y = y$ then $E[X|Y] = E[X|Y = y]$.

$E[X|Y] = g(Y)$

$E[X|Y=y] = \int_{S_X} x \, f_{X|Y}(x|y) \, dx$

- E[X|Y] returns E[X|Y=y] when Y=y
- Note that E[X|Y] is a random variable
  - It is a function of the RV ?
- E[X|Y=y] is a function of *y*
- E[E[X|Y]] is the expectation of a function of RV Y
- E[E[X|Y=1]] = ?
  - E[X|Y=1] ?

$\int_{y \in S_Y} \int_{S_X} x \, f_{X|Y}(x|y) \, dx \, f_Y(y) \, dy$

$f_{X,Y}(x,y) \, dx \, dy$

VVVS Problem

$$E[X|Y=y] = \int_y^1 x \frac{1}{1-y}\,dx$$

$$= \frac{1}{1-y}\left(\frac{1-y^2}{2}\right)$$

$$= \frac{1+y}{2}$$

Example 4.20 Problem

For random variables X and Y in Example 4.5, we found in Example 4.19 that the conditional PDF of X given Y is

$$E[X|Y] = \frac{1+Y}{2}$$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.109}$$

Find the conditional expected values $E[X|Y=y]$ and $E[X|Y]$.

- E[X|Y=y] = (1+y)/2
- E[X|Y] = (1+Y)/2

Independent Random Variables

Definition 4.16 *Independent Random Variables*

Random variables X *and* Y *are* independent *if and only if*

Discrete: $$P_{X,Y}(x, y) = P_X(x)P_Y(y)$$

Continuous: $$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

- For independent X and Y, $P_{X|Y}(x|y)$ = ?
  - $P_X(x)$
- What about $F_{X,Y}(x,y)$?

$$\frac{P_{X,Y}(x,y)}{P_Y(y)} = \frac{P_X(x)P_Y(y)}{P_Y(y)} = P_X(x)$$

$$P[X \le x, Y \le y] = P[X \le x]P[Y \le y]$$

Independent RVs

Theorem 4.27

For independent random variables X and Y,

(a) $E[g(X)h(Y)] =$,

(b) $r_{X,Y} = E[XY] =$ $E[X]E[Y]$

(c) $\text{Cov}[X, Y] =$ 0 ,

(d) $\text{Var}[X + Y] =$ $\text{Var}[X] + \text{Var}[Y]$,

(e) $E[X|Y = y] =$ $E[X]$,

(f) $E[Y|X = x] =$.

$$E[g(X)h(Y)]$$

$$= \int_{S_Y}\int_{S_X} g(x)h(y)\frac{f_{X,Y}(x,y)}{dx\,dy}$$

$$= \int_{S_Y}\int_{S_X} g(x)h(y)\, f_X(x) f_Y(y)\, dx\,dy$$

$$= E[g(X)]E[h(Y)]$$

Independent RVs

Theorem 4.27

For independent random variables X and Y,

(a) $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$,

(b) $r_{X,Y} = E[XY] = E[X]E[Y]$,

(c) $\text{Cov}[X, Y] = \rho_{X,Y} = 0$,

(d) $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$,

(e) $E[X|Y = y] = E[X]$ for all $y \in S_Y$,

(f) $E[Y|X = x] = E[Y]$ for all $x \in S_X$.

Example 4.24 Problem

$$f_{U,V}(u,v) = \begin{cases} 24uv & u \geq 0, v \geq 0, u+v \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.133}$$

Are U and V independent?

$f_U(u)$ $f_V(v)$

Quiz 4.10(A)

$P[X=x]\ P[Y=y]$ $\quad$ $P[X=x, Y=y]$

Random variables X and Y in Example 4.1 and random variables Q and G in Quiz 4.2 have joint PMFs:

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

$P_X(x) \longrightarrow P[X=1] = 0.18$

$P[X=2] = 0.81$

$P_Y(y)$

$P[Y=0] = 0.1$ $\quad$ $P[X=0] = 0.01$

$P[Y=1] = 0.09$

$P[Y=2] = 0.81$

(1) Are X and Y independent?

(2) Are Q and G independent?

Quiz 4.10(B)

iid

Random variables X_1 and X_2 are independentand identically distributed with probability density function

$$f_X(x) = \begin{cases} 1 - x/2 & 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.144}$$

$$\int_{-\infty}^{x_2}\int_{-\infty}^{x_1} f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2$$

$0 \le x_1 \le 2$

(1) What is the joint PDF $f_{X_1,X_2}(x_1, x_2)$?

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \left(1-\frac{x_1}{2}\right)\left(1-\frac{x_2}{2}\right) & 0 \le x_1 \le 2,\ 0 \le x_2 \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(2) Find the CDF of $Z = \max(X_1, X_2)$.

$$P[Z \le z] = P[\max(X_1,X_2) \le z] \to P[X_1 \le z, X_2 \le z]$$

Problem 4.10.13

X and Y are independent random variables with PDFs

$$\frac{P[x \le X < x+\Delta x \mid X>Y]}{\Delta x}$$

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let $A = \{X > Y\}$. $\{X>y\}$ $\{X>Y\}$

$f_{X|A}(x)$

(a) What are $E[X]$ and $E[Y]$?

(b) What are $E[X|A]$ and $E[Y|A]$?

$$f_{X|\{X>0.5\}} = \begin{cases} \dfrac{f_X(x)}{P(X>0.5)} & 0.5 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X|\{X>Y\}}(x) = \int f_{X,Y|A}(x,y)\,dy = \begin{cases} \dfrac{f_{X,Y}(x,y)}{P(X>Y)} & (x,y) \in A,\ x>y \quad 0 \le y \le x \le 1 \\ 0 \end{cases}$$

$f_{X,Y|A}(x,y)$

A Shorter Exponential

Let $X \sim Exp(\lambda)$ and $Y \sim Exp(\mu)$. Assume that X and Y are independent random variables. Find $f_{X|X<Y}(x)$.

HW

Extra Slides

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Problem 4.9.14

Suppose you arrive at a bus stop at time 0 and at the end of each minute, with probability p, a bus arrives, or with probability $1 - p$, no bus arrives. Whenever a bus arrives, you board that bus with probability q and depart. Let T equal the number of minutes you stand at a bus stop. Let N be the number of buses that arrive while you wait at the bus stop.

(a) Identify the set of points (n, t) for which $P[N = n, T = t] > 0$.

(b) Find $P_{N,T}(n, t)$.

(c) Find the marginal PMFs $P_N(n)$ and $P_T(t)$.

(d) Find the conditional PMFs $P_{N|T}(n|t)$ and $P_{T|N}(t|n)$.

$[p(1-q)]^{t-1} pq$

Problem 4.9.15

Each millisecond at a telephone switch, a call independently arrives with probability p. Each call is either a data call (d) with probability q or a voice call (v). Each data call is a fax call with probability r. Let N equal the number of milliseconds required to observe the first 100 fax calls. Let T equal the number of milliseconds you observe the switch waiting for the first fax call. Find the marginal PMF $P_T(t)$ and the conditional PMF $P_{N|T}(n|t)$. Lastly, find the conditional PMF $P_{T|N}(t|n)$.

Bivariate Gaussian Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Joint Density of X and Y is Gaussian

Bivariate Gaussian Random Variables

X, Y $\quad E[X]$ $\quad E[Y]$

$Var[X]$ $\quad Cov[X,Y]$

$Cov[X,Y]$ $\quad Var[Y]$

2×2

Definition 4.17 *Random variables X and Y have a* bivariate Gaussian PDF *with parameters μ_1, σ_1, μ_2, σ_2, and ρ if*

$$f_{X,Y}(x,y) = \frac{\exp\left[-\dfrac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

where μ_1 and μ_2 can be any real numbers, $\sigma_1 > 0$, $\sigma_2 > 0$, and $-1 < \rho < 1$.

- We say that the RV(s) X and Y are $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

Examples of a Joint Gaussian PDF

Figure 4.5

The Joint Gaussian PDF $f_{X,Y}(x, y)$ for $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and three values of ρ.

If X and Y are Jointly Gaussian then they Marginally Gaussian

Theorem 4.28

If X and Y are the bivariate Gaussian random variables in Definition 4.17, X is the Gaussian (μ_1, σ_1) random variable and Y is the Gaussian (μ_2, σ_2) random variable:

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} \quad f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-\mu_2)^2/2\sigma_2^2}.$$

- **How do we prove this?**
- Marginally Gaussian may not imply jointly Gaussian

For Jointly Gaussian RVs the Conditional PDF is Gaussian

Theorem 4.29

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

$$\tilde{\mu}_2(x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \tilde{\sigma}_2^2 = \sigma_2^2(1 - \rho^2).$$

E[Y|X=x]

Var[Y|X=x]

- E[Y|X] is a line that passes through (μ_1,μ_2). Also called regression line for any X and Y.
- (x, E[Y|X=x]) is also the point which maximizes $f_{Y|X}(y|x)$
- Following the values of $f_{X,Y}$ for points along the line implies following the maxima of all $f_{Y|X}(y|x)$

The Conditional Density

Figure 4.6

Each bell shaped curve corresponds to $f_{Y|X}(y|x)$ scaled by $f_X(x)$

Example Joint Gaussian Density

Correlation Coefficient

Theorem 4.31

Bivariate Gaussian random variables X and Y in Definition 4.17 have correlation coefficient

$$\rho_{X,Y} = \rho.$$

- We want to show that $\rho = \rho_{X,Y} = Cov[XY]/(\sigma_X\ \sigma_Y) = Cov[XY]/(\sigma_1\ \sigma_2)$
- $Cov[XY] = E[(X - \mu_X)\ (Y - \mu_Y)] = E[(X - \mu_1)\ (Y - \mu_2)]$
- Using Iterated Expectation we get
  - $E[(X - \mu_1)\ (Y - \mu_2)] = E[(X - \mu_1)\ E[(Y - \mu_2)\,|\,X]]$
- Theorem 4.29 gave us $E[Y|X]$

Correlation Coefficient

- We have

$$E[Y-\mu_2|X=x] = E[Y|X=x]-\mu_2$$
$$= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)-\mu_2$$

- Therefore $E[(X-\mu_1)\,E[(Y-\mu_2)|X]]$ can be written as

$$E[(X-\mu_1)\rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)] = \rho\frac{\sigma_2}{\sigma_1}E[(X-\mu_1)^2]$$

- Therefore, $\rho_{X,Y} = Cov[XY]/(\sigma_1\,\sigma_2) = \rho$

Uncorrelated and Independent

Theorem 4.32

Bivariate Gaussian random variables X and Y are uncorrelated if and only if they are independent.

- If they are independent, we know that they are uncorrelated (Covariance is 0)
- Given that X and Y are jointly Gaussian and that Cov[X,Y] = 0, we have ρ=0
- How do we show that X and Y are independent?

Uncorrelated => Independent for Jointly Gaussian

We know that the joint density is given by

Random variables X *and* Y *have a* bivariate Gaussian PDF *with parameters* μ_1, σ_1, μ_2, σ_2, *and* ρ *if*

$$f_{X,Y}(x,y) = \frac{\exp\left[-\dfrac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

where μ_1 *and* μ_2 *can be any real numbers,* $\sigma_1 > 0$, $\sigma_2 > 0$, *and* $-1 < \rho < 1$.

Substituting $\rho = 0$, we get $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Example When RVs are Uncorrelated but not Independent (Also see link)

- Consider two coin tosses. Let the outcome of the first toss be denoted by the RV X ~ Bernoulli(0.5) and the outcome of the second toss be modeled as RV Y ~ Bernoulli(0.5). Let the outcomes be independent of each other

- Clearly Cov[X,Y] = 0.
- Let Z = X – Y and W = X + Y, $S_Z = \{-1,0,1\}$ and $S_W = \{0,1,2\}$
- P[Z=0] = ½
- P[Z=0|W=2] = P[X=Y|W=2] = 1
- P[Z=1] = ¼ and P[Z=1|W=2] = 0
- Clearly W and Z are not independent
- However Cov[W,Z] = E[(X – Y) (X+Y-2μ)] = 0
- W and Z are uncorrelated but NOT independent

Variance of the Conditional Density

- The variance of the conditional density was defined in **Theorem 4.29**

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

$$\tilde{\mu}_2(x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \tilde{\sigma}_2^2 = \sigma_2^2(1 - \rho^2).$$

- Var[Y|X] <= Var[Y], since ρ is the corr coeff and hence|ρ| <= 1
- If Y and X are correlated, knowing X is likely to reduce the *uncertainty* in Y

Buffon's Needle!

A fine needle of length $2a$ is dropped at random on a board covered with parallel lines distance $2b$ apart where $b > a$ as in figure. What is the probability that the needle intersects one of the lines?

Linear Combinations of Normal RV(s)

1. Let $Z = aX + bY$, where X and Y are **jointly normal**. Then Z is normally distributed.
   - How will you show this?
   - Find $P[Z <= z]$, then $f_Z(z)$ and use the jointly normal distribution. Do this for at least when X and Y are independent
2. Let $Z = aX + bY$ and $W = cX + dY$. X and Y are **jointly normal.** Then Z and W are jointly normal
   - Proof requires additional machinery
   - Note that 1 follows from 2

Probability and Random Processes

Sums of RV

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Expected Value of Sum is Sum of Expected Values

Theorem 6.1

For any set of random variables $X_1, \ldots, X_n$, the expected value of $W_n = X_1 + \cdots + X_n$ is

$$E[W_n] = E[X_1] + E[X_2] + \cdots + E[X_n].$$

- Proof by induction
- We earlier proved that $E[W_2] = E[X_1] + E[X_2]$
- Assume $E[W_{n-1}] = E[X_1] + E[X_2] + \ldots + E[X_{n-1}]$
- Show $E[W_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$
  - $E[W_n] = E[W_{n-1} + X_n] = E[W_{n-1}] + E[X_n]$ (from the base case). Hence proved

Expected Value of Sum is Sum of Expected Values

- The theorem is valid for any joint distribution of the random variables $X_1, X_2, \ldots, X_n$
- The variables do not need to be independent! Can have ANY joint distribution
- Therefore a very powerful theorem!
- Can you show why the above is true using the chain rule for a pdf?
- The chain rule states that

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = f_{X_2,\ldots,X_n|X_1}(x_2,\ldots,x_n|x_1)f_{X_1}(x_1)$$
$$= f_{X_n|X_{n-1}\ldots X_1}(x_n|x_{n-1},\ldots,x_1)f_{X_{n-1}|X_{n-2}\ldots X_1}(x_{n-1}|x_{n-2},\ldots,x_1)\ldots f_{X_1}(x_1)$$

Variance of a Sum Of Random Variables

Theorem 6.2

The variance of $W_n = X_1 + \cdots + X_n$ is

$$\mathrm{Var}[W_n] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{Cov}\left[X_i, X_j\right].$$

$$\mathrm{Cov}[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])]$$

$$E[X_i X_j] - E[X_i]E[X_j]$$

- This is also the sum of the elements of the Covariance Matrix of the random variables X_1, X_2, …, X_n

$$E[X_i X_j] = \int_{x_j \in S_{X_j}} \int_{x_i \in S_{X_i}} x_i x_j \, f_{X_i X_j}(x_i, x_j) \, dx_i \, dx_j$$

$$f_{X_i}(x_i) f_{X_j}(x_j)$$

Theorem 6.3

If X, Y are independent then $E[XY] = E[X]E[Y]$

When $X_1, \ldots, X_n$ are uncorrelated,

$$\mathrm{Var}[W_n] = \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n].$$

~~If $E[XY] = E[X]E[Y]$ then X and Y are independent~~

$x_1, x_2, \ldots, x_n$

$$\frac{x_1 + x_2 + \cdots + x_n}{n}$$

$W_n =$

$$\frac{X_1 + X_2 + \cdots + X_n}{n}$$

0.5

0.5, 0.5

1 0

PDF of Sum of Two Random Variables

$X \le w - Y$

Theorem 6.4

$P[W \le w] = P[X + Y \le w]$

The PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_{X,Y}(x, w - x)\, dx = \int_{-\infty}^{\infty} \boxed{f_{X,Y}(w - y, y)}\, dy.$$

$$\int_{-\infty}^{\infty} \frac{d}{dw} \int_{-\infty}^{w-y} f_{X,Y}(u,v)\, du\, dv$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{w-y} f_{X,Y}(u,v)\, du\, dv$$

$$\int_{-\infty}^{\infty} f_{X,Y}(w-y, v)\, dv$$

Theorem 6.5

When X and Y are independent random variables, the PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_X(w - y)\, f_Y(y)\, dy = \int_{-\infty}^{\infty} f_X(x)\, f_Y(w - x)\, dx.$$

Moment Generating Function

Moment Generating Function

Definition 6.1 *(MGF)*

$$\frac{d}{ds}\phi_X(s) = \frac{d}{ds}\int_{-\infty}^{\infty} e^{sX} f_X(x)\,dx$$

$$= \int_{-\infty}^{\infty} \frac{d}{ds} e^{sX} f_X(x)\,dx$$

For a random variable X, the moment generating function (MGF) *of X is*

$$\phi_X(s) = E\left[e^{sX}\right]. \longrightarrow \int_{-\infty}^{\infty} e^{sX} f_X(x)\,dx$$

- What about the continuous case?

$$\frac{d}{ds}\phi_X(s) = \int_{-\infty}^{\infty} X e^{sX} f_X(x)\,dx$$

- One-to-one mapping between MGF of a RV and its PDF
  - See table in textbook

Calculating Moments Using the MGF

Theorem 6.6

A random variable X with MGF $\phi_X(s)$ has nth moment

$$\underline{E\left[X^n\right]} = \left.\frac{d^n \phi_X(s)}{ds^n}\right|_{s=0}.$$

Example 6.5 Problem

X is an exponential random variable with MGF $\phi_X(s) = \lambda/(\lambda - s)$. What are the first and second moments of X? Write a general expression for the nth moment.

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0.$$

$$\phi_X(s) = E\left[e^{sX}\right] = \int_0^\infty e^{sx} \lambda e^{-\lambda x} dx$$

MGF of a Scaled and Shifted RV

Theorem 6.7

The MGF of $Y = aX + b$ is $\phi_Y(s) = e^{sb}\phi_X(as)$.

$E[e^{sY}]$

Theorem 6.8: MGF of Sum of Independent Random Variables

For a set of independent random variables $X_1, \ldots, X_n$, the moment generating function of $W = X_1 + \ldots + X_n$ is

$$E[e^{sW}] = E[e^{s(X_1 + X_2 + \cdots + X_n)}]$$

$$= E[e^{sX_1} e^{sX_2} \cdots e^{sX_n}]$$

$$= E[e^{sX_1}] \cdots E[e^{sX_n}]$$

$$\phi_W(s) = \phi_{X_1}(s)\phi_{X_2}(s)\ldots\phi_{X_n}(s).$$

When $X_1, \ldots, X_n$ are iid, each with MGF $\phi_{X_i}(s) = [\phi_X(s)]^n$,

$$\phi_W(s) = (\phi_X(s))^n.$$

Example 6.6 Problem

J and K are independent random variables with probability mass functions

$$P_J(j) = \begin{cases} 0.2 & j = 1, \\ 0.6 & j = 2, \\ 0.2 & j = 3, \\ 0 & \text{otherwise,} \end{cases} \qquad P_K(k) = \begin{cases} 0.5 & k = -1, \\ 0.5 & k = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.40}$$

Find the MGF of $M = J + K$? What are $E[M^3]$ and $P_M(m)$?

- Note that we are looking at discrete random variables
- Therefore

$$\phi_M(s) = \sum_{m_i \in S_M} e^{s m_i} P_M(m_i)$$

Sum of Independent Poisson Random Variables is Poisson Distributed

Theorem 6.9

If $K_1, \ldots, K_n$ are independent Poisson random variables, $W = K_1 + \cdots + K_n$ is a Poisson random variable.

- The PMF of Poisson RV is

$$P_{K_i}(x) = \begin{cases} \alpha_i^x e^{-\alpha_i}/x! & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

- The MGF is

$$\phi_{K_i}(s) = e^{(\alpha_i)(e^s - 1)}$$

- Now we use Theorem 6.8

Sum of Independent Gaussian Random Variables is Gaussian Distributed

Theorem 6.10

The sum of n independent Gaussian random variables $W = X_1 + \cdots + X_n$ is a Gaussian random variable.

- The MGF of a Gaussian is

$$\phi_{X_i}(s) = e^{s\mu_i + s^2\sigma_i^2/2}$$

- Can you show the above?

Sum of Independent Exponentials

Theorem 6.11

If $X_1, \ldots, X_n$ are iid exponential (λ) random variables, then $W = X_1 + \cdots + X_n$ has the Erlang PDF

$$f_W(w) = \begin{cases} \frac{\lambda^n w^{n-1} e^{-\lambda w}}{(n-1)!} & w \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

- The MGF of an exponential is

$$\phi_{X_i}(s) = \left(\frac{\lambda}{\lambda - s}\right)$$

Quiz 6.4(B)

Let $X_1, \ldots, X_n$ be independent Gaussian random variables with $E[X_i = 0]$ and $\mathrm{Var}[X_i] = i$. Find the PDF of

$$W = \alpha X_1 + \alpha^2 X_2 + \cdots + \alpha^n X_n. \tag{6.54}$$

e^{sW}

Problem 6.8.4

In a subway station, there are exactly enough customers on the platform to fill three trains. The arrival time of the nth train is $X_1 + \cdots + X_n$ where $X_1, X_2, \ldots$ are iid exponential random variables with $E[X_i] = 2$ minutes. Let W equal the time required to serve the waiting customers.

MGF of W $\longleftrightarrow$ CDF of W

Find $P[W > 20]$.

Central Limit Theorem

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Central Limit Theorem

Figure 6.1

Binomial: Sum of which RV(s)?

Bernoulli (p) iid

$\frac{W_n}{\sqrt{n}} \rightarrow \frac{W_n - n\mu_X}{\sqrt{n}\sigma_X} \xrightarrow{n\to\infty}$

$n = 5$ $n = 10$ $n = 20$

$n\mu_X$

$\sqrt{n}\,\sigma_X^2$

$Var[X_1 + \cdots + X_n] = n\sigma_X^2$

The PMF of the X, the number of heads in n coin flips for $n = 5, 10, 20$. As n increases, the PMF more closely resembles a bell-shaped curve.

- Which PMF are we talking about?

Mean, Variance

$W_n = X_1 + X_2 + \cdots + X_n$

$X_i \sim X_j$ $E[X] = \mu_X$, $Var[X] = \sigma_X^2$

Central Limit Theorem

Theorem 6.14 Central Limit Theorem

Given $X_1, X_2, \ldots$, a sequence of iid random variables with expected value μ_X and variance σ_X^2, the CDF of $Z_n = (\sum_{i=1}^{n} X_i - n\mu_X)/\sqrt{n\sigma_X^2}$ has the property

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z).$$

- Equivalently, we can write

$$\frac{S_n - nE[X]}{\sqrt{n\sigma_X^2}} \xrightarrow{D} \Phi(z) \text{ as } n \to \infty.$$

- That is Z_n *converges in distribution* to the standard normal distribution

Figure 6.2

(a) $n = 1$ **(b)** $n = 2$

(c) $n = 3$, **(d)** $n = 4$

The PDF of W_n, the sum of n uniform $(0, 1)$ random variables, and the corresponding central limit theorem approximation for $n = 1, 2, 3, 4$. The solid — line denotes the PDF $f_{W_n}(w)$, while the $- \cdot -$ line denotes the Gaussian approximation.

Note that here n=4 is large enough!

Figure 6.3

The binomial (n, p) CDF and the corresponding central limit theorem approximation for $n = 4, 8, 16, 32$, and $p = 1/2$.

Definition 6.2 Approximation

Let $W_n = X_1 + \cdots + X_n$ *be the sum of* n *iid random variables, each with* $E[X] = \mu_X$ *and* $\mathrm{Var}[X] = \sigma_X^2$*. The central limit theorem approximation to the CDF of* W_n *is*

$$F_{W_n}(w) \approx \Phi\left(\frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right).$$

Quiz 6.6

The random variable X milliseconds is the total access time (waiting time + read time) to get one block of information from a computer disk. X is uniformly distributed between 0 and 12 milliseconds. Before performing a certain task, the computer must access 12 different blocks of information from the disk. (Access times for different blocks are independent of one another.) The total access time for all the information is a random variable A milliseconds.

(1) What is $E[X]$, the expected value of the access time?

(2) What is $\mathrm{Var}[X]$, the variance of the access time?

(3) What is $E[A]$, the expected value of the total access time? $\rightarrow 12 \times 6$

(4) What is σ_A, the standard deviation of the total access time? $\rightarrow n\left(\frac{12^2}{12}\right) = 12n$

(5) Use the central limit theorem to estimate $P[A > 75 \text{ ms}]$, the probability that the total access time exceeds 75 ms. $P[A - 72 > 3] = P\left[\frac{A-72}{\sqrt{144}} > \frac{3}{\sqrt{144}}\right]$

(6) Use the central limit theorem to estimate $P[A < 48 \text{ ms}]$, the probability that the total access time is less than 48 ms.

$P[\text{Bit is } 1] = p$

$P[499000 \le \text{No. of ones} \le 501000]$

$= P[\text{No of ones} \le 501000]$

$- P[\quad]$

Example 6.15 Problem

Transmit one million bits. Let A denote the event that there are at least 499,000 ones but no more than 501,000 ones. What is $P[A]$?

Definition 6.3 De Moivre–Laplace Formula

For a binomial (n, p) *random variable* K,

$$P\left[k_1 \le K \le k_2\right] \approx \Phi\left(\frac{k_2 + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 0.5 - np}{\sqrt{np(1-p)}}\right).$$

Quiz 6.7

Telephone calls can be classified as voice (V) if someone is speaking or data (D) if there is a modem or fax transmission. Based on a lot of observations taken by the telephone company, we have the following probability model: $P[V] = 3/4$, $P[D] = 1/4$. Data calls and voice calls occur independently of one another. The random variable K_n is the number of voice calls in a collection of n phone calls.

(1) What is $E[K_{48}]$, the expected number of voice calls in a set of 48 calls?

(2) What is $\sigma_{K_{48}}$, the standard deviation of the number of voice calls in a set of 48 calls?

(3) Use the central limit theorem to estimate $P[30 \leq K_{48} \leq 42]$, the probability of between 30 and 42 voice calls in a set of 48 calls.

(4) Use the De Moivre–Laplace formula to estimate $P[30 \leq K_{48} \leq 42]$.