3. General Random Variables

Han-You Jeong

School of Electrical and Computer Engineering

Pusan National University

Probability and Statistics

Outline

Probability Density Function (PDF)

- A random variable X is called continuous if there is a nonnegative function f_X, called the probability density function (PDF) of X such that

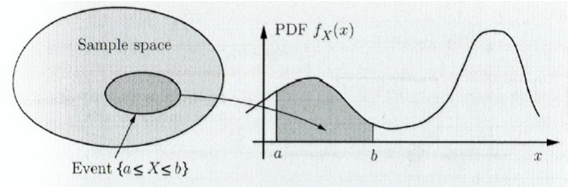
$$P(X \in B) = \int_B f_X(x)dx,$$

for every subset B of the real line.

  - The probability that the value of X falls within an interval $[a, b]$ is

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx, \text{ and } P(a \leq X \leq a) = \int_a^a f_X(x)dx = 0.$$

  - $P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$.

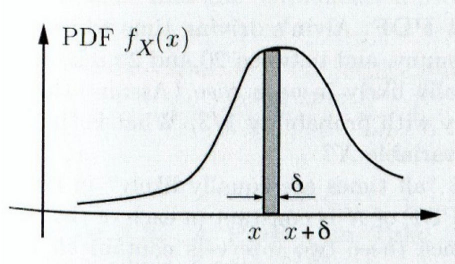


- From the normalization property, we have

$$\int_{-\infty}^{\infty} f_X(x)dx = P(-\infty < X < \infty) = 1.$$

- For an interval $[x, x+\delta]$ with very small length δ, we have

$$P([x, x+\delta]) = \int_{x}^{x+\delta} f_X(t)dt \approx f_X(x) \cdot \delta,$$



- $f_X(x)$ is neither the probability of any particular event nor restricted to be less than or equal to one.

A Couple of Examples

Ex 3.2 Alvin's driving time to work is between 15 and 20 minutes if the day is sunny, and between 20 and 25 minutes if the day is rainy, with all times being equally likely in each case. Assume that a day is sunny with probability $2/3$ and rainy with probability $1/3$. What is the PDF of the driving time?

Ex 3.3 Consider a random variable X with PDF

$$f_X(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & \text{if } 0 < x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Is it a valid PDF?

Expectation of a Continuous Random Variables

- The expectation of a continuous random variable X with PDF f_X is

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

- The expected value rule for a function $g(X)$ has the form

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

- The variance of X is defined by

$$var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x) dx.$$

- We have $0 \leq var(X) = E[X^2] - (E[X])^2$.
- If $Y = aX + b$, where a and b are given scalars, then

$$E[Y] = aE[X] + b, \quad var(Y) = a^2 var(X).$$

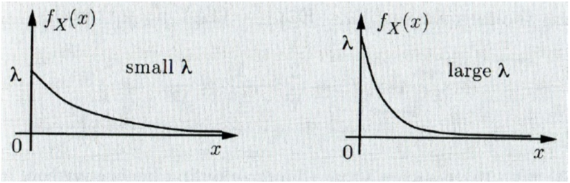
Ex 3.4 Calculate the mean and the variance of uniform random variable over the interval $[a, b]$.

Exponential Random Variable

- An exponential random variable has a PDF of the form

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

where λ is a positive parameter characterizing the PDF.



- An exponential random variable is a good model for the amount of time until an incident takes place.
  - a message arriving at a computer, some equipment breaking down, a light bulb burning, etc.
- Show that $E[X] = 1/\lambda$ and $var(X) = 1/\lambda^2$.

Cumulative Distribution Functions (CDFs)

- The cumulative distribution function (CDF) of a random variable X is denoted by $F_X(x)$ and provides the probability $P(X \leq x)$.

$$F_X(x) = P(X \leq x) = \begin{cases} \displaystyle\sum_{k \leq x} p_X(k), & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{x} f_X(t)dt, & \text{if } X \text{ is continuous.} \end{cases}$$

  - The CDF $F_X(x)$ accumulates probability up to the value x.
  - $F_X(x)$ is monotonically nondecreasing, i.e., if $x \leq y$, then $F_X(x) \leq F_Y(y)$.
  - $F_X(x)$ tends to 0 as $x \to -\infty$, and to 1 as $x \to \infty$.
  - If X is discrete, then $F_X(x)$ is a piecewise constant function of x.
  - If X is continuous, then $F_X(x)$ is a continuous function of x.

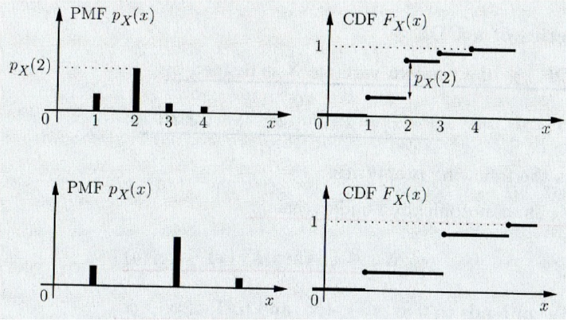
CDFs of Discrete Random Variables

- If X is discrete and take integer values, we have

$$F_X(k) = \sum_{i=-\infty}^{k} p_X(i),$$

$$p_X(k) = P(X \le k) - P(X \le k-1) = F_X(k) - F_X(k-1),$$

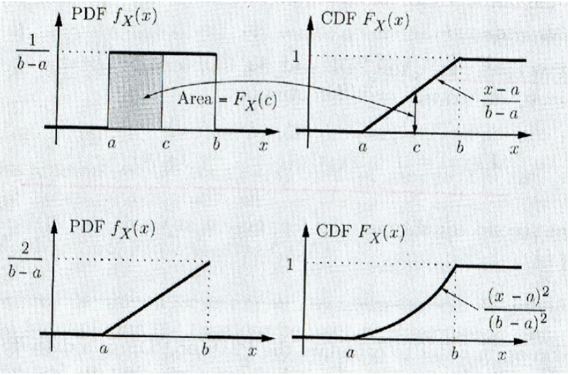
for all integers k.



CDFs of Continuous Random Variables

- If X is continuous, we have

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \text{ and } f_X(x) = \frac{dF_x(x)}{dx}.$$



The Geometric and Exponential CDFs

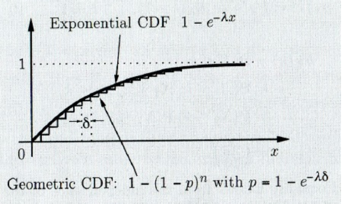
- The CDF provides a convenient means for exploring the relations between continuous and discrete random variables.
  - Let X be a geometric random variable with parameter p, then

$$P(X=k)=p(1-p)^{k-1}, \text{ and } F_{geo}(n)=\sum_{k=1}^{n}P(X=k)=1-(1-p)^n.$$

  - Suppose that X be an exponential random variable with parameter λ,

$$P(X=x)=\lambda e^{-\lambda x}, \text{ and } F_{exp}(x)=\begin{cases}0, & \text{for } x\le 0,\\ 1-e^{-\lambda x}, & \text{for } x>0.\end{cases}$$

  - By setting $e^{-\lambda\delta}=1-p$, we have $F_{exp}(n\delta)=F_{geo}(n)$.



Normal Random Variables

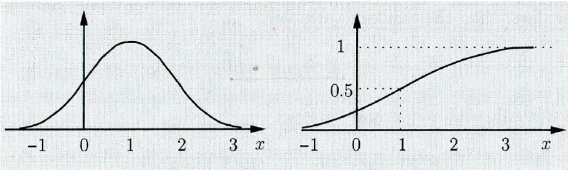
- A continuous random variable X is said to be normal or Gaussian if it has a PDF of the form

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

where μ is the mean and $\sigma > 0$ is the standard deviation.

- From the normalization property, we have

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



Normal PDF $f_X(x)$ Normal CDF $F_X(x)$

The Standard Normal Random Variable

- If a random variable Y is a linear transform of random variable X, i.e. $Y = aX + b$ ($a \neq 0$), it is also normal with mean and variance

$$E[Y] = a\mu + b, \quad var(Y) = a^2\sigma^2.$$

- A normal random variable Y with zero mean and unit variance is said to be a standard normal with CDF Φ

$$\Phi(y) = P(Y \leq y) = P(Y < y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt.$$

  - Using the symmetry of the PDF, we have $\Phi(-y) = 1 - \Phi(y)$ for all y.
- We standardize normal random variable X with mean μ and variance σ^2 by defining a new random variable Y given by

$$Y = \frac{X - \mu}{\sigma}, \text{ then } E[Y] = \frac{E[X] - \mu}{\sigma} = 0 \text{ and } var(Y) = \frac{var(X)}{\sigma^2} = 1.$$

The Standard Normal Table

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Central Limit Theorem

- A normal random variable models well additive effect of many independent factors in a variety of engineering, physical, and statistical contexts.
- Central Limit Theorem
  - The sum of a large number of independent and identically distributed (i.i.d.) random variables has an approximately normal CDF regardless of the CDF of the individual random variables.

Ex 3.7 The annual snowfall in Busan is modeled as a normal random variable with a mean of $\mu = 60$ mm and a standard deviation of $\sigma = 20$ mm. What is the probability that this year's snowfall will be at least 80 mm

Ex 3.8 A binary message is transmitted as a signal s, which is either -1 or $+1$. The communication channel corrupts the transmission with additive normal noise with mean $\mu = 0$ and variance σ^2. What is the probability of bit error?

Joint PDFs

- Two continuous random variables associated with the same experiment are said to be jointly continuous and can be described by a joint PDF $f_{X,Y}$, where

$$P((X,Y)\in B)=\iint\limits_{(x,y)\in B} f_{X,Y}(x,y)dxdy,$$

  for every subset B of the two-dimensional plane.

- From the normalization property, we have

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

- Let δ be a small positive number, we have in the vicinity of (a,c):

$$P(a\le X\le a+\delta, c\le Y\le c+\delta)=\int_{c}^{c+\delta}\int_{a}^{a+\delta} f_{X,Y}(x,y)dxdy\approx f_{X,Y}(a,c)\delta^2$$

Marginal PDFs

- The joint PDF can be used to calculate the probability of an event involving only one of them. For example, let A be a subset of the real line and consider the event $\{X \in A\}$

$$P(X \in A) = \int_A f_X(x)dx = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx.$$

- From the above, the marginal PDFs f_X and f_Y are given by

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

Ex 3.11 A surface is ruled with parallel lines, which are at distance d from each other. Suppose that we throw a needle of length l on the surface at random. What is the probability that the needle will intersect one of the lines?

Joint CDFs and Expectations

- If X and Y are two random variables associated with the same experiment, their joint CDF is defined as

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)dt\ ds.$$

- Conversely the joint PDF can be recovered from the joint CDF:

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

- From the expected value rule, we have

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

- For any scalars a, b, and c, $E[aX + bY + c] = aE[X] + bE[Y] + c$.

Ex 3.12 Let X and Y be described by a uniform PDF on the unit square. Find the joint PDF and the joint CDF.

More than Two Random Variables

- The joint PDF of three random variables X, Y, and Z is defined in analogy with the case of two random variables:

$$P((X,Y,Z)\in B)=\iiint_{(x,y,z)\in B} f_{X,Y,Z}(x,y,z)\; dxdydz.$$

- The marginal PDFs are given by

$$f_{X,Y}(x,y)=\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,z)dz, \text{ and } f_X(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}(x,y,$$

- From the expected value rule, we have

$$E[g(X,Y,Z)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y,z)f_{X,Y,Z}(x,y,z)dxdydz.$$

- For any random variables, $X_1, X_2, \cdots, X_n$ and any scalars $a_1, a_2, \cdots, a_n$, we have

$$E[a_1X_1+a_2X_2+\cdots+a_nX_n]=a_1E[X_1]+a_2E[X_2]+\cdots+a_nE[X_n].$$

Conditioning a Random Variable on an Event

- The conditional PDF of a continuous random variable X, given an event A with $P(A) > 0$, is defined as a nonnegative function $f_{X|A}$ that satisfies

$$P(X \in B|A) = \int_B f_{X|A}(x)dx,$$

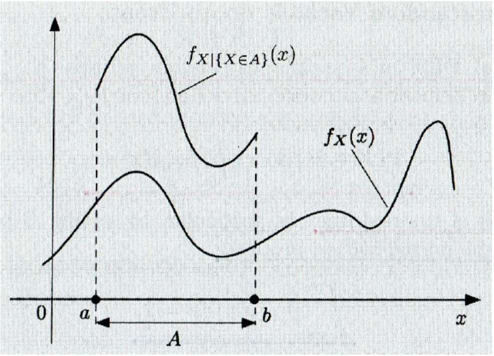
for any subset B of the real line.

- From the normalization property

$$\int_{-\infty}^{\infty} f_{X|A}(x)dx = 1.$$

- The conditional PDF has the same shape as the unconditional one scaled by the constant factor $1/P(X \in A)$ so that $f_{X|A}$ integrates to 1

$$f_{X|A}(x) = \begin{cases} \dfrac{f_X(x)}{P(X \in A)}, & \text{if } x \in A, \\ 0, & \text{otherwise.} \end{cases}$$



Ex 3.13 The time T until a new light bulb burns out is an exponential random variable with parameter λ. Ariadne turns the lights on , leaves the room, and when she returns, t times units later, finds that the light bulb is still on, i.e., $A = \{T > t\}$. Let X be the additional time until the light bulb burns out. What is the conditional CDF of X, given the event A?

- Let X and Y be jointly continuous random variable with joint PDF $f_{X,Y}$. If we condition on an event $C = \{(X, Y) \in A\}$ with $P(C) > 0$,

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

- The conditional PDF of X, given this event is

$$f_{X|C}(x) = \int_{-\infty}^{\infty} f_{X,Y|C}(x, y)dy.$$

- If the events $A_1, \cdots, A_n$ form a partition of the sample space, then

$$f_X(x) = \sum_{i=1}^{n} P(A_i) f_{X|A_i}(x).$$

Ex 3.14 The metro train arrives at the station every quarter hour starting at 6:00 a.m. You walk into the station between 7:10 and 7:30 a.m., and your arrival time is a uniform random variable over this interval. What is the PDF of the time you have to wait for the first train to arrive?

Conditioning one Random Variable on Another

- For any y with $f_Y(y) > 0$, the conditional PDF of random variable X given that $Y = y$, is defined by

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

- Furthermore, the formula

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

implies that

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

Ex 3.15 Ben throws a dart at a circular target of radius r. We assume that he always hits the target and that all points of impact (x, y) are equally likely. Find the joint PDF $f_{X,Y}(x, y)$, the marginal PDF $f_Y(y)$, and the conditional PDF $f_{X|Y}(x|y)$.

- For small positive numbers δ_1 and δ_2, we have

$$P(x \le X \le x+\delta_1 | y \le Y \le y+\delta_2) \approx \frac{f_{X,Y}(x,y)\delta_1\delta_2}{f_Y(y)\delta_2} = f_{X|Y}(x|y)\delta_1,$$

and more generally,

$$P(X \in A | Y = y) = \int_A f_{X|Y}(x|y)dx.$$

- For the case of more than two random variables, we have

$$f_{X,Y|Z}(x,y|z) = \frac{f_{X,Y,Z}(x,y,z)}{f_Z(z)}, \text{ and } f_{X|Y,Z}(x|y,z) = \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)}.$$

- From an analog of the multiplication rule,

$$f_{X,Y,Z}(x,y,z) = f_{X|Y,Z}(x|y,z)f_{Y|Z}(y|z)f_Z(z).$$

Ex 3.16 The speed of a vehicle driving past a police radar is modeled as an exponential R. V. X with mean 50 Km/h. The police radar's measurement Y of the speed has an error modeled as a normal R. V. with zero mean and standard deviation equal to one tenth of the speed. Find the joint PDF $f_{X,Y}(x,y)$.

Conditional Expectation

- Let X and Y be jointly continuous R. V.s, and let A be an event with $P(A) > 0$.
- The conditional expectations are given by

$$E[X|A] = \int_{-\infty}^{\infty} x f_{X|A}(x)dx, \text{ and } E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx.$$

- From the expected value rule, we have

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

- From the total expectation theorem, we have

$$E[g(X, Y)] = \int E[g(X, Y)|Y = y] f_Y(y)dy.$$

Ex 3.17 Suppose that the R. V. X has the piecewise constant PDF, where $f_X(x) = 1/3$ for $0 \leq x \leq 1$; $f_X(x) = 2/3$ for $1 < x \leq 2$; otherwise, $f_X(x) = 0$. Find $E[X]$ and $var(X)$.

Independence

- Two continuous R. V.s X and Y are independent if their joint PDF is the product of the marginal PDFs:

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

- The independence is the same as the condition

$$f_{X|Y}(x|y) = f_X(x), (f_Y(y) > 0), \text{ and } f_{Y|X}(y|x) = f_Y(y)(f_X(x) > 0).$$

- The independence implies that

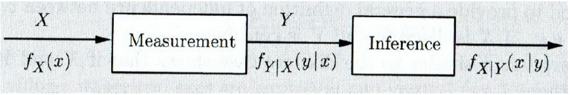
$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) = F_X(x)F_Y(y).$$

- If X and Y are independent, then $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$.

Ex 3.18 Let X and Y be independent normal R. V.s with means μ_x, μ_y, and variances σ_x^2, σ_y^2, respectively. Find the joint PDF and the contour curve.

Basic Idea of Continuous Bayes' Rule

- We usually represent an unobserved phenomenon by a R. V. X with PDF f_X and we make a noisy measurement Y, which is modeled in terms of a conditional PDF $f_{Y|X}$.



- Here, we are interested in the conditional PDF $f_{X|Y}$, i.e.,

$$f_{X|Y}(x|y) = \frac{f_X(x)f_{Y|X}(y|x)}{f_Y(y)} = \frac{f_X(x)f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(t)f_{Y|X}(y|t)dt}.$$

Ex 3.19 The lifetime of a light bulb Y follows an exponential distribution. On any given day, the parameter λ of the PDF Y is also a R. V. uniformly distributed in $[1, 3/2]$. We test a light bulb and record its lifetime. What can we say about the underlying parameter λ?

Inference about a Discrete Random Variable

- In some cases, the unobserved phenomenon is inherently discrete.
- Let p_N be the PMF of N and Y be a continuous R. V. whose conditional PDF, given value n of N, is described by $f_{Y|N}(y|n)$. Then, we have

$$P(N = n|Y = y) = \frac{p_N(n) f_{Y|N}(y|n)}{f_Y(y)} = \frac{p_N(n) f_{Y|N}(y|n)}{\sum_{\forall i} p_N(i) f_{Y|N}(y|i)}.$$

Ex 3.20 A binary signal S is transmitted, and we are given that $P(S = 1) = p$ and $P(S = -1) = 1 - p$. The received signal is $Y = N + S$, where N is normal noise with zero mean and unit variance, independent of S. What is the probability that $S = 1$, as a function of the observed value y of Y?

Inference based on Discrete Observations

- Let $P(A)$ be the probability of event A and Y be a continuous R. V. and assume that the conditional PDFs $f_{Y|A}(y)$ and $f_{Y|A^c}(y)$ are known.

$$P(A|Y=y) \approx \frac{P(A)f_{Y|A}(y)}{f_Y(y)} = \frac{P(A)f_{Y|A}(y)}{P(A)f_{Y|A}(y) + P(A^c)f_{Y|A^c}(y)}.$$

- Given an event A, we make an inference about a R. V. Y as follows:

$$f_{Y|A}(y) = \frac{f_Y(y)P(A|Y=y)}{P(A)} = \frac{f_Y(y)P(A|Y=y)}{\int_{-\infty}^{\infty} f_Y(t)P(A|Y=t)dt}.$$

Homework # 3

- Section 3.1 Continuous Random Variables and PDFs: Problem 1
- Section 3.2 Cumulative Distribution Functions: Problem 7
- Section 3.3 Normal Random Variables: Problem 13
- Section 3.4 Joint PDFs of Multiple Random Variables: Example 3.11 Problem 15
- Section 3.5 Conditioning: Examples 3.13 Problems 22 and 25
- Section 3.6 The Continuous Bayes' Rule: Example 3.19 Problem 34