COMP 233/2

Probability and Statistics for Computer Science

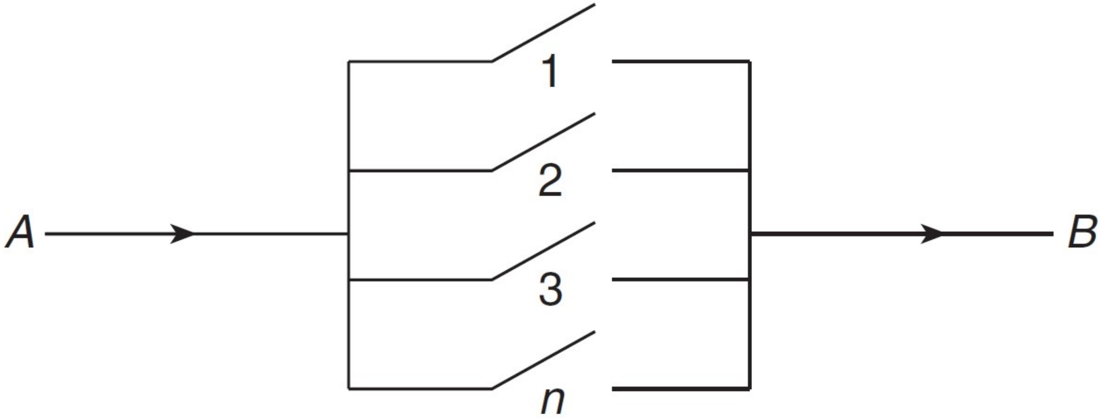
Week 3 (slides curtsey Dr. T. Fevens)

S. P. Mudur & H. Mellah

Random variable
Probability mass function
Probability density function
Cumulative distribution function
Joint probability mass function

Reading: Chap 4

Example



- *A* system is composed of *n* separate components.
- It is ***parallel*** if the system functions when at least one component does.
- It is also assumed that the operation of one component does not affect the other ones.
- Suppose that the probability that the *k*-th component functions is p_k. Thus the probability that the *k*-th component fails is $1 - p_k$.
- *Find the probability that the system functions.*

Solution

- Let A be the event that the system functions.
- The system is *parallel*, hence

$$A = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n = \bigcup_{k=1}^{n} A_k$$

where A_k is the event that the k-th component functions.

Solution

- By De Morgan Laws, $A^c = \left(\bigcup_{k=1}^{n} A_k \right)^c = A_1^c A_2^c A_3^c \cdots A_n^c$
- Independence implies that

$$P(A^c) = \prod_{k=1}^{n} P(A_k^c) = \prod_{k=1}^{n} (1 - p_k)$$

- Hence,

$$P(A) = 1 - P(A^c) = 1 - \prod_{k=1}^{n} (1 - p_k)$$

Probability as a Function

- So far the probability value that we have associated with any event has been an explicit number (a numerical value).
- We have relied on the assumption that every outcome of an experiment is equally likely.
- Hence we can use the formula: $P(E) = \dfrac{N(E)}{N(S)}$
- What do we do when this is not the case?
- **Probability is defined as a function (discrete or continuous) of one or more variables.**

Example 1

- When a bit is transmitted, it is received in error with probability 0.1.
- Assume that 8 bits are transmitted independently.
  - How many bits are likely to be received in error?
  - What is the probability that 2 bits are received in error?

Solution

- Let "e" and "O" stand for a bit received in error and an "ok" bit, respectively.
- Consider the following two outcomes.

e	O	O	O	e	O	O	O
e	e	O	O	O	O	O	e

- There are 2 errors in the first byte and 3 errors in the second one.

Solution

- Independence implies that

$$
\begin{aligned}
&P(\mathrm{eOeOOOOO}) \\
&= P(E_1 O_2 E_3 O_4 \ldots O_8) \\
&= \mathrm{P}(E_1)\mathrm{P}(O_2)\mathrm{P}(E_3)\mathrm{P}(O_4)\ldots\mathrm{P}(O_8) = 0.1^2 \cdot 0.9^6.
\end{aligned}
$$

- Similarly, $\mathrm{P}(\mathrm{eeOOOOOe}) = 0.1^3 \cdot 0.9^5.$

Solution

- Equally likely, however, are *all outcomes with exactly two* erroneous bits.
- It remains to answer a question:

  how many different outcomes with two "e"'s are there?

1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}
O	e	e	O	O	O	O	O

Solution

- In other words, in how many ways can we select 2 bits out of 8, that are received in error?
- The answer is provided by the respective *binomial* coefficient $\binom{8}{2}$.
- Hence, the requested probability is $\binom{8}{2} 0.1^2 0.9^6$.

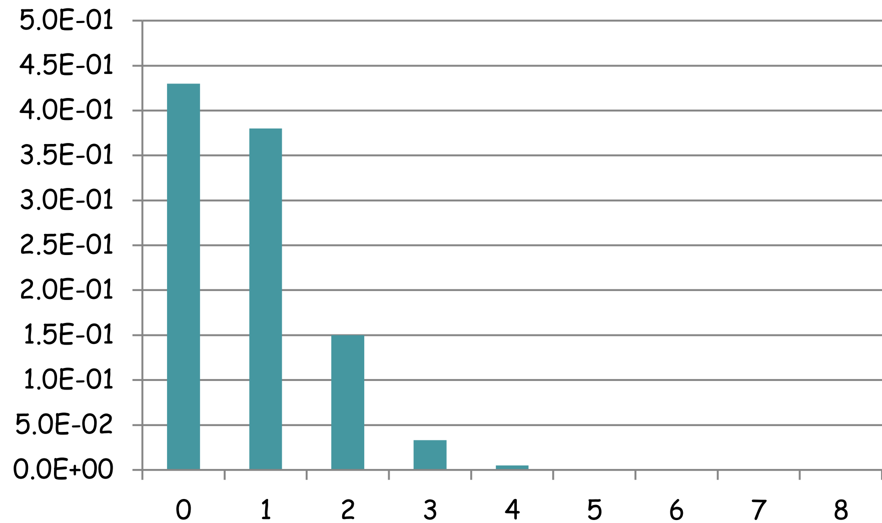
1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}
O	e	e	O	O	O	O	O

Solution

- To summarize, let X stand for the number of bits received in error.
- Then

$$P\{X = \mathrm{k}\} = \binom{8}{k} 0.1^{\mathrm{k}} \cdot 0.9^{8-\mathrm{k}}$$

where k is an integer between 0 and 8.



k	0	1	2	3	4	5	6	7	8
P{X=k}	0.43	0.38	0.15	0.033	0.005	0.0004	2.27 E-05	7.2 E-07	1.0 E-08

General Case

- In general, let X stand for the number of bits received in error, when *n* bits are transmitted, with the probability of a single bit in error being *p*.
- Then

$$P\{X = \mathrm{k}\} = \binom{n}{k} p^{\mathrm{k}} (1-p)^{n-\mathrm{k}}$$

where *k* is an integer between 0 and *n*.

Observation

- We recall that

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

- Hence,

$$\sum_{k=0}^{n} P\{X = k\} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= (p+1-p)^n = 1$$

Example 2

- When a bit is transmitted, it is received in error with probability 0.1.
- Assume that bits are transmitted independently, until the first bit is received in error.
  - How many bits can be received?
  - What is the probability that the 5-th bit is received in error?
  - What is the probability that at least 5 (i.e. 5 or more) bits are received until the first error?

Solution

- Experiment: Transmitting until the first error.
  - S = {e, 0e, 00e, 000e, 0000e, … }
  - So infinitely many bits can be received.

  - The probability that the 5-th bit is received in error is
  $$P(0000e) = (0.1)(0.9)^4$$

  - Let X stand for the number of bits received until the first error.

Solution

- Therefore, the probability that the k-th bit is first bit received in error is

$$P\{X = \mathrm{k}\} = 0.1 \cdot 0.9^{\mathrm{k}\text{-}1}$$

where k is a positive integer. Note that $k \geq 1$ since at least one bit must be received.

k	1	2	3	4	5	6	7	8	9	…
P{X=k}	**0.100**	**0.090**	**0.081**	**0.073**	**0.066**	**0.059**	**0.053**	**0.048**	**0.043**	**…**

General Case

- In general, let X stand for the number of bits received until the first error, with the probability of a single bit in error being *p*.
- Then

$$P\{X = \mathrm{k}\} = p(1 - p)^{\mathrm{k}-1}$$

where *k* is a positive integer.

Geometric series

- Let -1 < q < 1
- Then the series $$G(q) = \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$
- is convergent (absolutely and uniformly for every open sub interval of (-1, 1).
- Note: $$\sum_{k=0}^{\infty} P\{X = k\} = p\sum_{k=0}^{\infty}(1-p)^k = 1.$$

Solution

- In order to answer the last question we need to evaluate an infinite sum.
- To find is

$$P\{X \geq 5\} = \sum_{k=5}^{\infty} P\{X = k\} = 0.1 \sum_{k=5}^{\infty} 0.9^{k-1}$$

- The last sum is a geometric series with $q = 0.9$.
- Hence, the sum equals

$$0.1\left[\sum_{k=1}^{\infty} 0.9^{k-1} - (0.9^0 + 0.9^1 + 0.9^2 + 0.9^3)\right]$$

$$= 0.1\left[\frac{1}{1-0.9} - (0.9^0 + 0.9^1 + 0.9^2 + 0.9^3)\right]$$

$$= 0.1[10 - 3.439] = 0.6561.$$

Summary

- In preceding examples, we started with *[non-numerical]* *event* *sample spaces*.
- We managed to associate *random numbers* with every outcome of respective sample space.
- In addition, for each *random number* we derived a function for the *probability* of any value occurring.
- We also observed that these *probabilities sum up to 1* in each example.

Random Variables

Random Variables

- *A **random variable*** (RV) X is a quantity of interest defined on the sample space of a probability experiment.
- That is, a RV is a value determined by the outcome of the experiment that assumes its possible values with certain probabilities.
- As with sample spaces, we can have discrete or continuous random variables.

Discrete Random Variables

- If a sample space *S* is *discrete*, then every RV defined on *S* is also ***discrete***, i.e. its range is countable.

Examples of Discrete Random Variables

- The outcome (1-6) of rolling a single die.
- The sum of the numbers on the dice, when k dice are rolled.
- The number of bits received until the r-th error.
- The number of bits received in error when n bits are received.

Example

- A random variable of possible interest for the "rolling two dice" experiment is the value of the their sum. Letting X denote this random variable, then X can take different values from 2 to 12.
- So,

$$P\{X = i\} = ???, \quad i = 2, 3, \ldots, 12$$

- Note that since X must take on some value,

$$P(S) = P\left(\bigcup_{i=2}^{12}\{X = i\}\right) = \sum_{i=2}^{12} P\{X = i\} = 1$$

Continuous Random Variables

- If a sample space S is *continuous*, then every RV defined on S is also ***continuous***, i.e. its set of possible values form an interval (a continuum of possible values).

Examples of Continuous Random Variables

- The time between consecutive phone calls.
- The diameter of a dot produced by a laser printer.
- The thickness of a wafer in semiconductor manufacturing.
- The current or voltage in an electric circuit.

Probability Mass Function (PMF)

- Using the associated probabilities $P\{X\}$ for values of a *discrete* RV X with values $x_1, x_2, \ldots$, we can easily construct its ***probability mass function***:

$$p(x) = \begin{cases} p_{\mathrm{i}}, \text{ if } x = x_{\mathrm{i}} \\ 0, \text{ otherwise} \end{cases}$$

where $p_i = P\{X = x_i\}$ and $\sum_{i \geq 1} p(x_i) = 1$

- Note that *p* is a function that is non-zero only on the range of X.

e	O	O	O	e	O	O	O

Probability Mass Function from Example 1

- The range is integer numbers from 0 to 8.
- The probabilities are given by

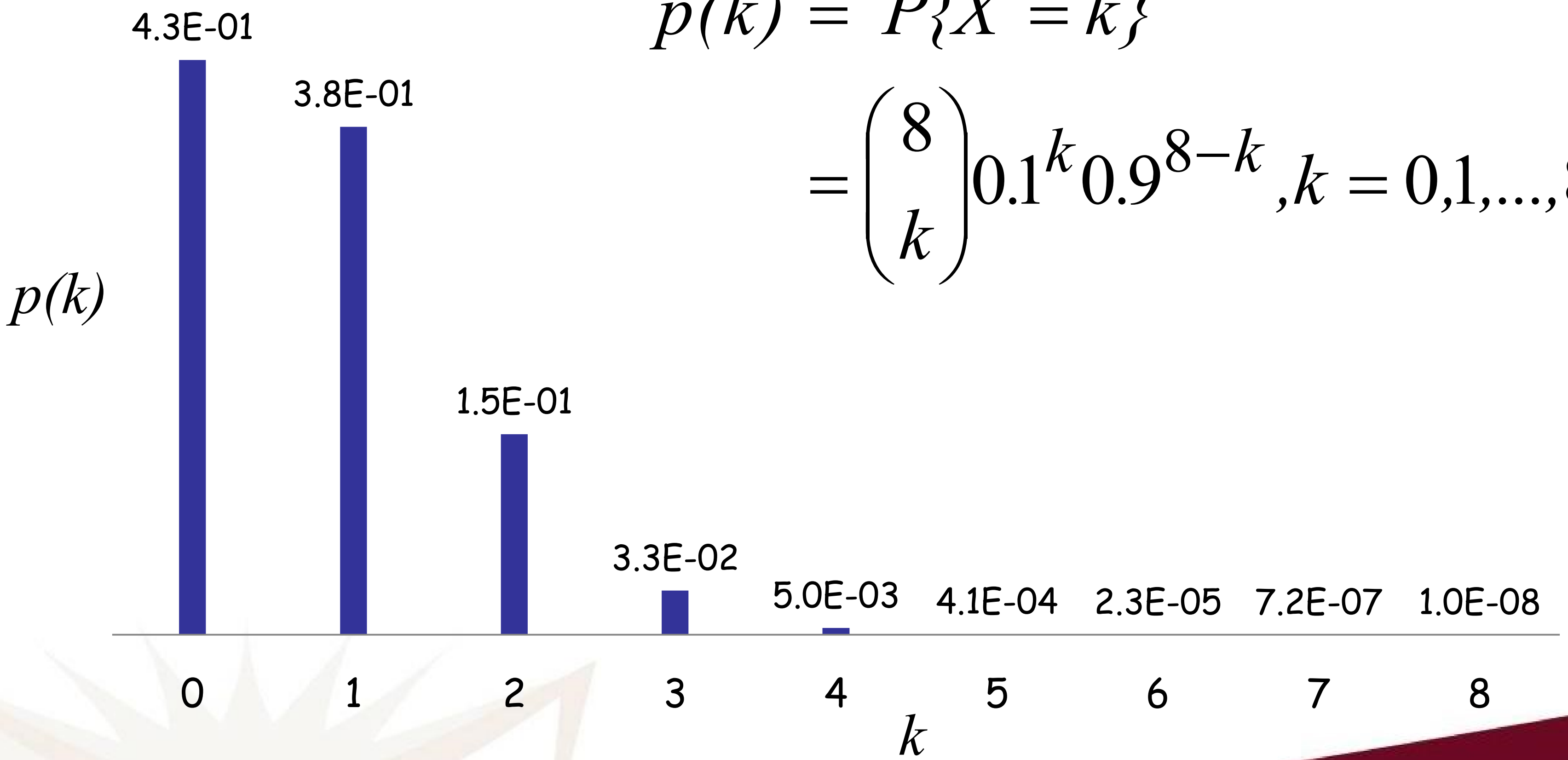
$$\mathrm{p(k)} = P\{X = \mathrm{k}\} = \binom{8}{k} 0.1^{\mathrm{k}} 0.9^{8-\mathrm{k}}, k = 0,1,\ldots,8$$

k	0	1	2	3	4	5	6	7	8
p(k)	0.43	0.38	0.15	0.033	0.005	0.00041	2.3E-05	7.2E-07	1E-08

PMF of Example 1

$$p(k) = P\{X = k\}$$

$$= \binom{8}{k} 0.1^k 0.9^{8-k}, k = 0,1,\ldots,8$$



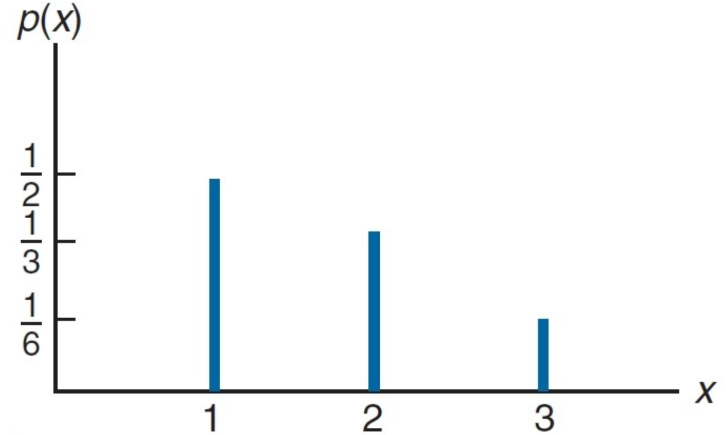
Cumulative Distribution Function (CDF)

- Often, we are interested in the probability of X ≤ a
- This is given by Cumulative Distribution Function (CDF)
- The CDF provides an alternative definition for a *discrete* RV X.

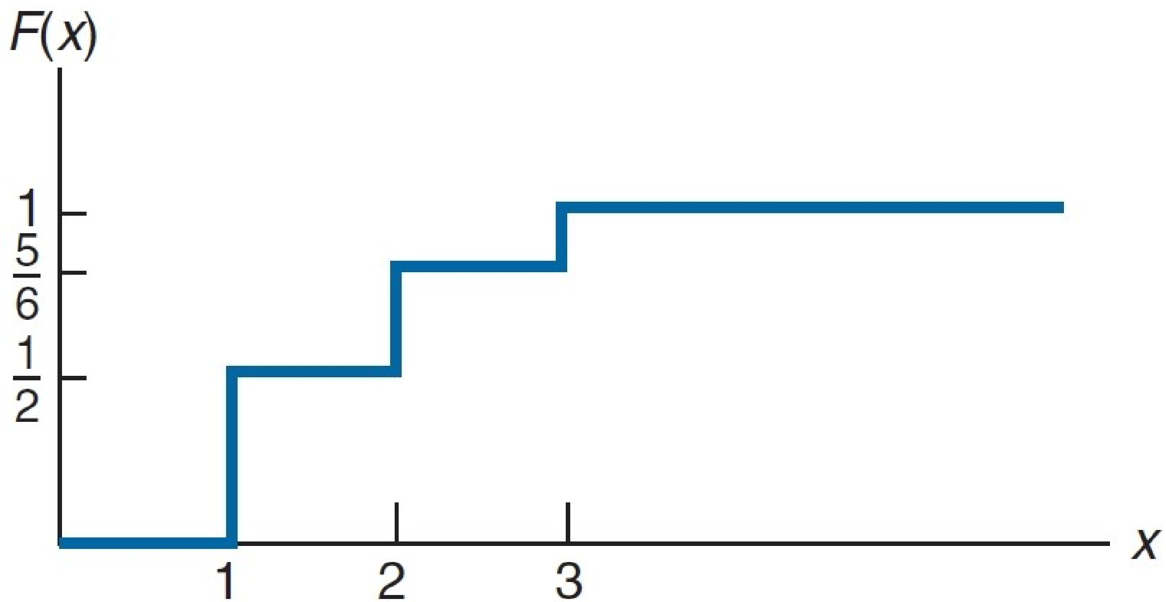
$$F(x) = P\{X \leq x\} = \begin{cases} 0, \text{ if } x < x_1 \\ P\{X = x_1\},\ x_1 \leq x < x_2 \\ P\{X = x_1\} + P\{X = x_2\}, x_2 \leq x < x_3 \\ \ldots \end{cases}$$

- We notice that *F* is a *step* function for a *discrete* RV .

Example



- Consider a random variable X that is equal to 1, 2, or 3. If we know that p(1) = 1/2, p(2) = 1/3, and p(3) = 1/6 (note that p(1) + p(2) + p(3) = 1) then its cumulative distribution function F(x) is shown below.



Cumulative Distribution Function

- The cumulative distribution function F can be expressed in terms of $p(x)$ by

$$F(a) = \sum_{\text{all } x \le a} p(x)$$

- So for Example 1, if $0 \le a \le 8$,

$$F(a) = \sum_{k=0}^{\lfloor a \rfloor} \binom{8}{k} 0.1^k 0.9^{8-k}$$

Probability Density Function

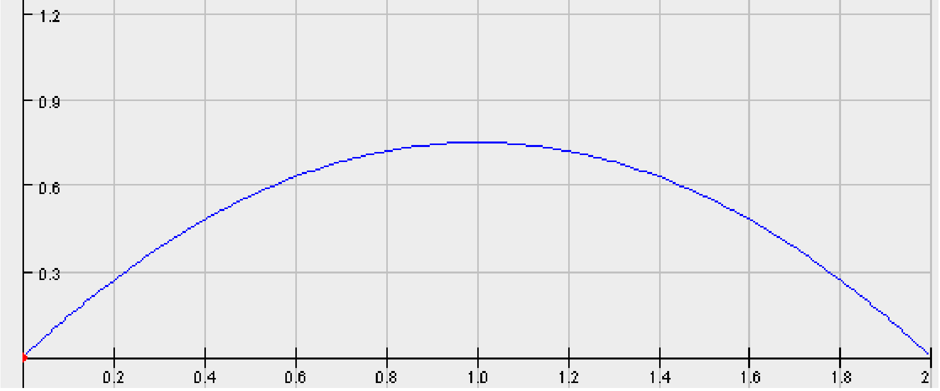
- For a ***continuous*** random variable X, the function f is called the ***probability density function*** (PDF), if

$$f(x) \geq 0,$$

$$x \in R, \text{ and}$$

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

Example



- An example of PDF assigned to a continuous random variable:

$$f(x) = \begin{cases} \frac{3}{8}(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

- Note that

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

Observation

- So far, we have been discussing problems in which the probability has been given to us (assuming it is based on previously collected statistical data) or we have been able to count the frequency and the size of the sample space.
- There are many situations in which the nature of the scientific scenario suggests a distribution type.
- When independent repeated observations are binary (dead/alive, success/failure, etc.) then it is the binomial distribution.
- The scenario of "time to failure" suggests an exponential distribution.
- Different Probability Distributions - Our topic for the next few weeks.

Example (exponential RV)

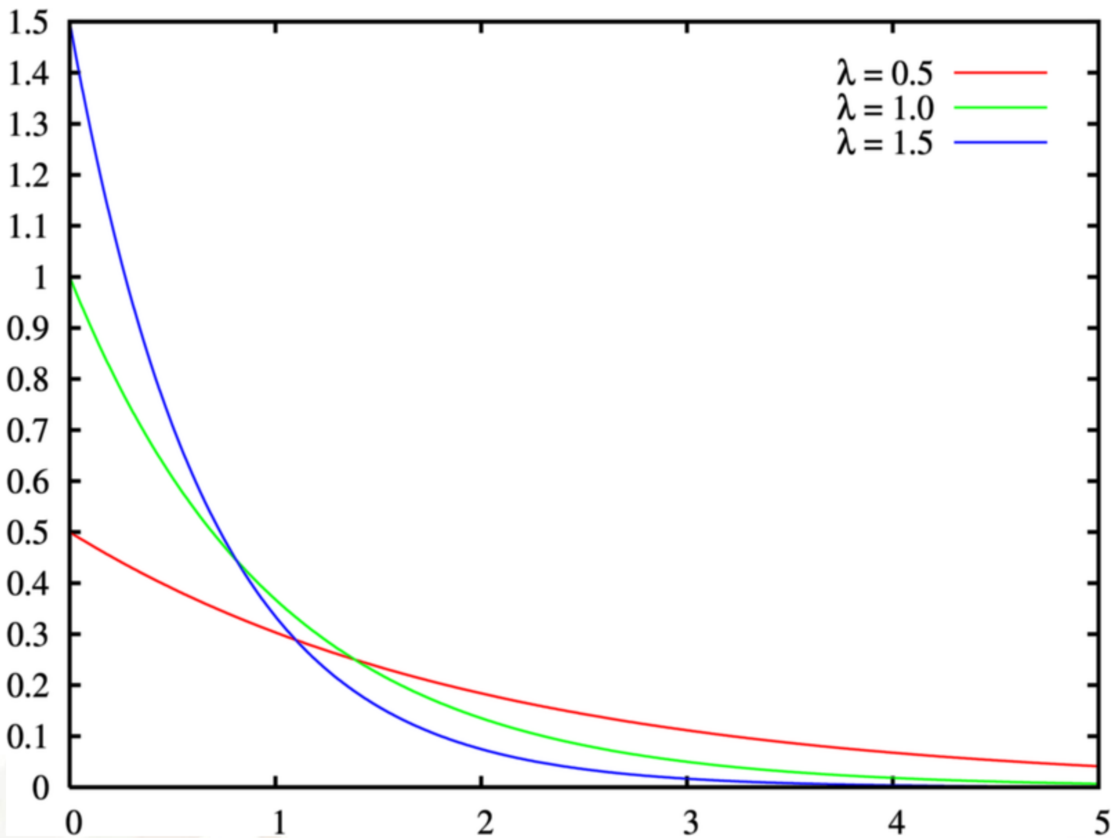
- The ***probability density function*** of the *exponential* RV with λ>0,

$$f(t) = \begin{cases} 0, t \le 0, \\ \lambda e^{-\lambda t}, t > 0, \end{cases}$$

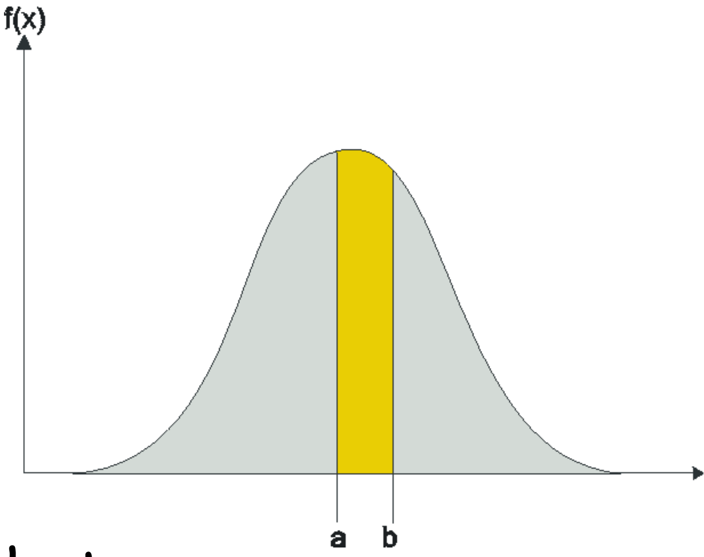
indeed, satisfies the definition:

- it is non-negative;
- the total area under its curve equals 1.

Probability Density Function (Exponential RV, various λ).



Probability



- *A continuous random variable* X is determined by its **probability density function**, $f(x)$, in the sense that

$$P\{a \le X \le b\} = \int_a^b f(x)dx.$$

- Important property:

$$P\{X = a\} = P\{a \le X \le a\} = \int_a^a f(x)dx = 0$$

Cumulative Distribution Function – Continuous RV

- By analogy with discrete mass functions, the ***cumulative distribution function*** for a continuous RV X is defined by

$$F(\mathrm{a}) = P\{X \leq \mathrm{a}\} = \int_{-\infty}^{\mathrm{a}} f(\mathrm{x})dx.$$

- We notice that F is non-negative and non-decreasing. In addition, $F(x)$→1 as x approaches infinity.

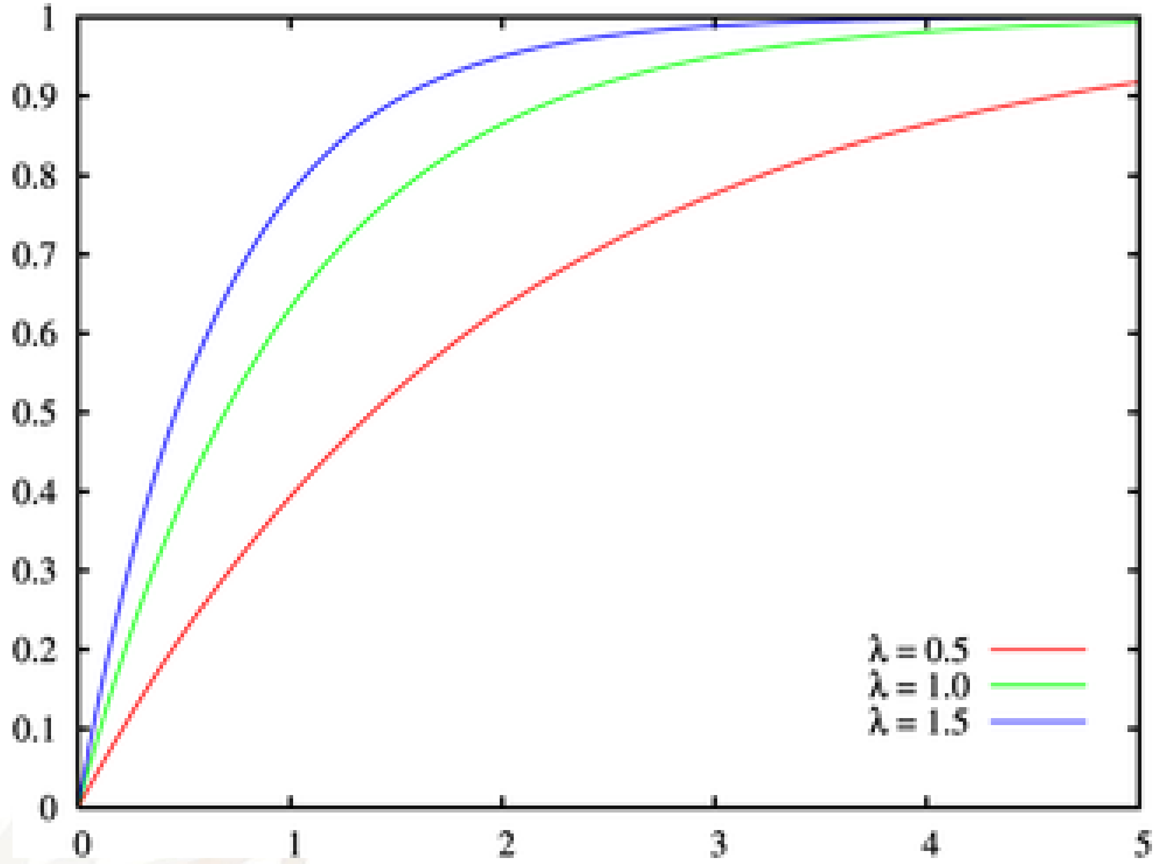
Cumulative Distribution Function – Example (exponential RV)

- The ***cumulative distribution function*** for the exponential RV is given by

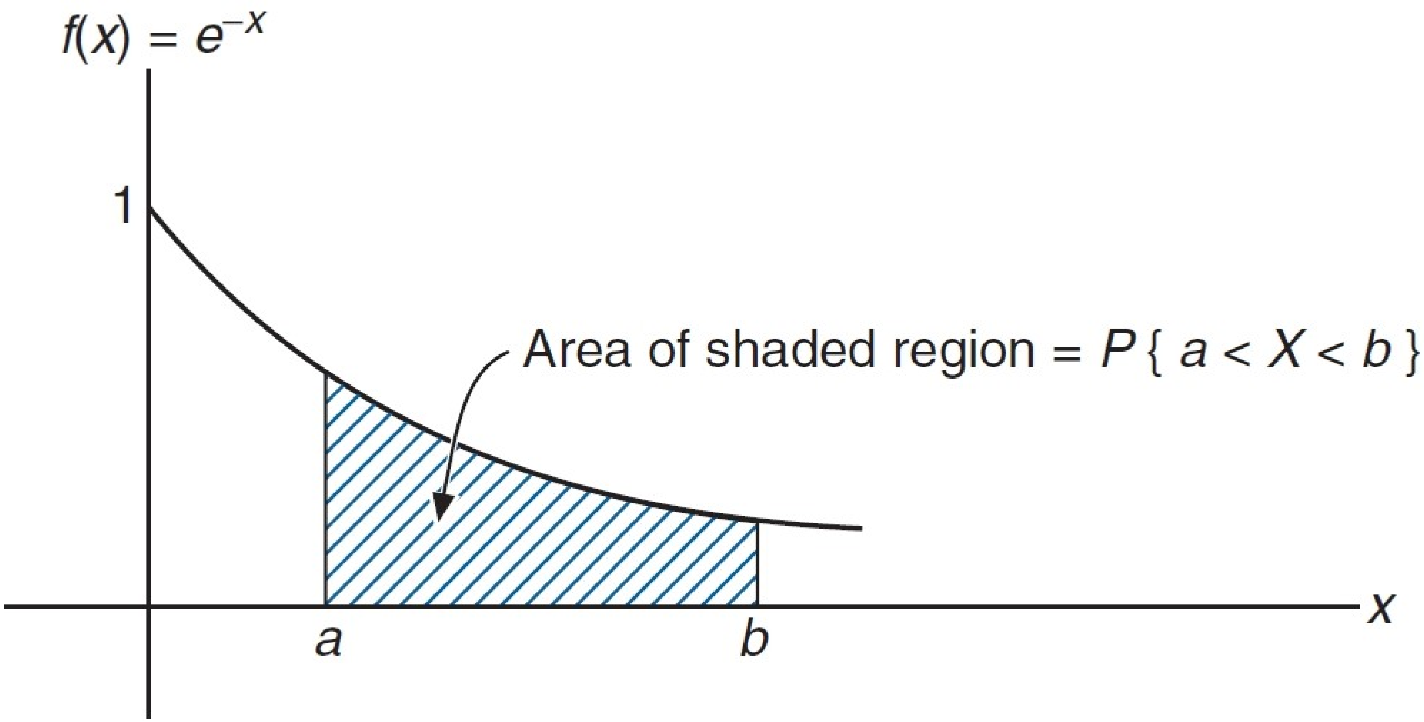
$$F(t) = P\{X \le t\} = \begin{cases} 0, \text{if } t \le 0, \\ 1 - e^{-\lambda t}, t > 0. \end{cases}$$

- As expected, we notice that F is non-negative and non-decreasing. Also, $F(t) \to 1$ as t approaches infinity.

CDF (Exponential RV, various λ)



Example (exponential RV)



$$\textit{The probability density function } f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

Deriving the PDF f(x) from F(x) in General

- Since

$$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(\mathrm{t})dt.$$

- It follows that

$$\frac{d}{dx}F(x) = f(x); \qquad P(a < X \le b) = F(b) - F(a);$$

- That is, the probability density function (PDF) is the derivative of the cumulative distribution function (CDF).

Example (exponential RV)

- The PDF $f(t)$ for the *exponential* RV can be determined from the CDF $F(t)$ by a "simple" derivative:

$$f(t) = \frac{d}{dt} F(t)$$

$$= \frac{d}{dt} \begin{cases} 0, & t \leq 0 \\ 1 - e^{-\lambda t}, & t > 0 \end{cases} \qquad = \begin{cases} 0, & t \leq 0 \\ \lambda e^{-\lambda t}, & t > 0 \end{cases}$$

Example

- Is the function

$$f(x) = \begin{cases} 0, & x < -1, \\ x^2, & -1 \le x \le 1, \\ 0, & x > 1, \end{cases}$$

a probability density function?

Example

- Is the function

$$f(\mathrm{x}) = \begin{cases} 0, & \mathrm{x} < -1, \\ 1.5\mathrm{x}^2, & -1 \le \mathrm{x} \le 1, \\ 0, & \mathrm{x} > 1, \end{cases}$$

a probability density function?

Example

- Find the value of c so that

$$f(\mathrm{x}) = \begin{cases} 0 \quad , & \mathrm{x} < 0, \\ c\mathrm{x}e^{-\mathrm{x}^2}, & \mathrm{x} \geq 0, \end{cases}$$

is a PDF.

Solution

- We observe that f is non-negative if c is so.
- It remains to ensure that the integral equals 1:

$$\int_{-\infty}^{\infty} f(\mathrm{x})d\mathrm{x} = 1 \Rightarrow \int_{0}^{\infty} c\mathrm{x}e^{-\mathrm{x}^2} d\mathrm{x} = 1$$

$$\Rightarrow \frac{c}{2}\int_{0}^{\infty} 2\mathrm{x}e^{-\mathrm{x}^2} d\mathrm{x} = 1 \Rightarrow \frac{c}{2}[-e^{-\mathrm{x}^2}]_0^{\infty} = 1$$

$$\Rightarrow \frac{c}{2} = 1 \Rightarrow c = 2.$$

Conclusion

- Hence, the function

$$f(\mathrm{x}) = \begin{cases} 0 \quad , & \mathrm{x} \le 0, \\ 2\mathrm{x}e^{-\mathrm{x}^2}, & \mathrm{x} > 0, \end{cases}$$

determines a continuous Probability Density Function.

Observation

- Similarly, the function

$$f(x) = \begin{cases} 0 \quad , & x \le 0, \\ 3x^2 e^{-x^3}, & x > 0, \end{cases}$$

determines a continuous Probability Density Function.

Weibull Density Function

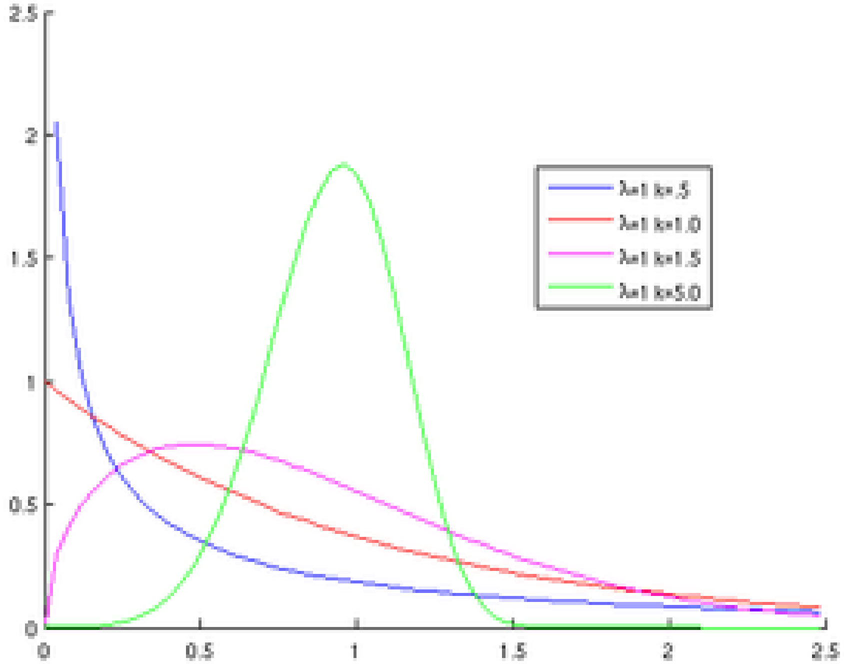
- The ***Weibull*** Density Function, with parameters $k > 0$ and $\lambda > 0$, is determined by its PDF:

$$f_{k,\lambda}(x) = \begin{cases} 0 \quad , & x \le 0, \\ \dfrac{k}{\lambda} \cdot \left(\dfrac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, & x > 0. \end{cases}$$

- k is called ***shape*** parameter, and λ is called ***scale*** parameter.
  - *Weibull* Density Function allows us to model <u>the time until failure</u> of various physical systems.

Weibull Density Function

- The PDF of *Weibull* Density Function with different $k > 0$ and $\lambda = 1$ are sketched below.



Uniform Density Function

- An RV is said to be ***uniformly*** distributed ***on*** *[α, β]* if its PDF is

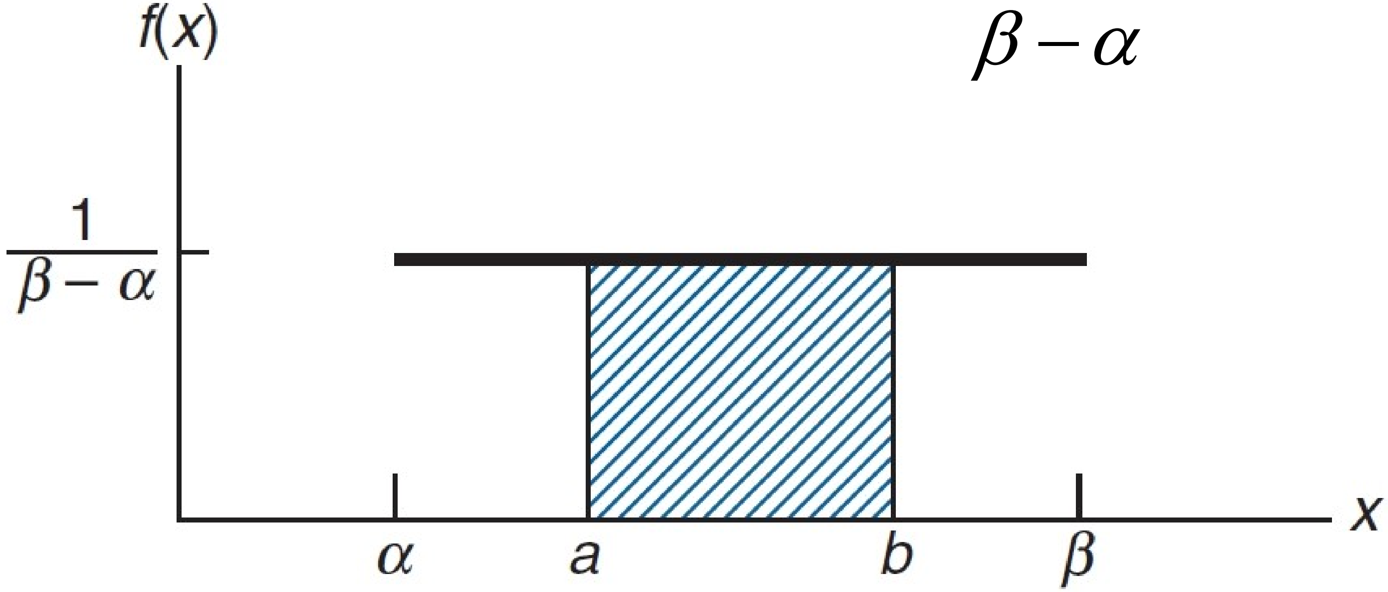
$$f(x) = \begin{cases} 0, & x < \alpha, \\ \dfrac{1}{\beta - \alpha}, & \alpha \le x \le \beta, \\ 0, & x > \beta. \end{cases}$$

  - The probability of any interval is only dependent on its *length*!
  - Under certain conditions, the time, the distance, the weight, etc., can be modeled using the uniform Density Function.

Uniform Density Function

$$P\{a \le X \le b\} = \frac{1}{\beta - \alpha}\int_a^b dx$$

$$= \frac{b-a}{\beta - \alpha}$$



CDF of Popular Density Functions

- Integration implies that

$$F(x) = 1 - e^{-\lambda x}, x > 0, \ (0, otherwise) \quad \text{Exponential;}$$

$$F(x) = 1 - e^{-(\frac{x}{\lambda})^k}, x > 0, \ (0, otherwise) \quad \text{Weibull;}$$

$$F(x) = \begin{cases} 0 \ , \ x \leq \alpha, \\ \dfrac{x-\alpha}{\beta-\alpha}, \alpha < x < \beta, \\ 1 \ , x \geq \beta. \end{cases} \quad \text{Uniform.}$$

  - CDFs are convenient for the calculation of the associated probabilities.

Jointly Distributed Random Variables

Observation

- Until this point we always dealt with one RV at a time, pdfs and cdfs in one dimension.
- In order to investigate probabilities depending on (two or more) RVs we need a tool that allows us to study these RVs in multiple dimensions at the same time.
- We saw that RV can be determined in terms of its Cumulative Distribution Function.
- We shall try extending this concept to involve more than one RV.

Joint Probability Mass Function

- If X and Y are **discrete** RV whose possible values are, respectively, x_1, x_2, …, and y_1, y_2, …, then

$$p(x_i, y_j) = P\{X = x_i, Y = y_j\}$$

- The function p is the ***joint probability mass function*** of X and Y.
- Needless to say, the sum of p over all x_i and y_j equals 1:

$$\sum_j \sum_i p(x_i, y_j) = 1.$$

Joint Probability Mass Function

- The event $\{X = x_i\}$ can be written as the union, over all j, of the <u>mutually exclusive</u> events $\{X = x_i,$ $Y = y_j\}$. That is, $\{X = x_i\} = \bigcup_j \{X = x_i, Y = y_j\}$
- Then, $P\{X = x_i\} = P\left(\bigcup_j \{X = x_i, Y = y_j\}\right)$

$$= \sum_j P\{X = x_i, Y = y_j\}$$

$$= \sum_j p(x_i, y_j)$$

Joint Cumulative Distribution Function

- We introduce the ***joint cumulative distribution function*** for RV X and Y by

$$F(x, y) = P\{X \leq x, Y \leq y\}.$$

- F is non-negative.
- In addition, $F(x,y) \rightarrow F(x)$, as y approaches infinity, and $F(x,y) \rightarrow F(y)$, as x approaches infinity.

Example

- Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries.
- If we let X denote the number of new batteries and Y the number of used but still working batteries that are chosen, then the joint probability mass function p of X and Y,

$$p(i,j) = P\{X = i, Y = j\},$$

is given by…

Example

$$p(0,0)=\binom{5}{3}\Big/\binom{12}{3}=\frac{10}{220}$$

$$p(0,1)=\binom{4}{1}\binom{5}{2}\Big/\binom{12}{3}=\frac{40}{220}$$

$$p(0,2)=\binom{4}{2}\binom{5}{1}\Big/\binom{12}{3}=\frac{30}{220}$$

$$p(0,3)=\binom{4}{3}\Big/\binom{12}{3}=\frac{4}{220}$$

$$p(1,0)=\binom{3}{1}\binom{5}{2}\Big/\binom{12}{3}=\frac{30}{220}$$

$$p(1,1)=\binom{3}{1}\binom{4}{1}\binom{5}{1}\Big/\binom{12}{3}=\frac{60}{220}$$

$$p(1,2)=\binom{3}{1}\binom{4}{2}\Big/\binom{12}{3}=\frac{18}{220}$$

$$p(2,0)=\binom{3}{2}\binom{5}{1}\Big/\binom{12}{3}=\frac{15}{220}$$

$$p(2,1)=\binom{3}{2}\binom{4}{1}\Big/\binom{12}{3}=\frac{12}{220}$$

$$p(3,0)=\binom{3}{3}\Big/\binom{12}{3}=\frac{1}{220}$$

Example

- Or in table form:

p(1,2)

used but still working

new

i \ j	0	1	2	3	P{X=i}
0	10/220	40/220	30/220	4/220	84/220
1	30/220	60/220	18/220	--	108/220
2	15/220	12/220	--	--	27/220
3	1/220	--	--	--	1/220
P{Y=j}	56/220	112/220	48/220	4/220	220/220

- The PMFs for X are obtained by the row sums, and the PMFs for Y are obtained by the column sums.

Example, cont.

- For the previous Joint Probability Mass Function for the batteries:
  a) Determine $F(2, 1)$.

  b) Determine $F_Y(1)$.

  c) What is the probability of 2 or more new batteries?

Joint Probability Density Function

- If the RV X and Y are **continuous**, then

$$F(s,t) = P\{X \le s, Y \le t\}$$

$$= \int_{-\infty}^{t} \int_{-\infty}^{s} f(x,y)\,dx\,dy.$$

- The function f is the ***joint probability density function*** of X and Y.
- Needless to say, the integration of f over the whole x-y plane equals 1 (F(∞,∞) = 1).

Development of Density for Sum.

- We shall compute the CDF of X from F(x,y).

$$
\begin{aligned}
F_X(t) &= P\{X < t\} \\
&= P\{X < t, -\infty < Y < \infty\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{t} f(x, y)\,dx\,dy \\
&= \int_{-\infty}^{t}\left(\int_{-\infty}^{\infty} f(x,y)\,dy\right)dx \\
&= \int_{-\infty}^{t} f_X(x)\,dx
\end{aligned}
$$

Example 1

- Let X stand for the time until a server connects to your machine, and let Y denote the time until the server authorizes you as a valid user. Each of the RV measures the wait from a common starting point. It is clear that $X < Y$.
- Under certain assumptions,

$$f(x, y) = \begin{cases} 0.000006\, e^{-0.001x - 0.002y} & 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

Example 1 (modified)

- Let X stand for the time until a server connects to your machine, and let Y denote the time from the moment of connection until the authorization as a valid user is complete. It is clear that now $X > 0$ and $Y > 0$.
- Under certain assumptions,

$$f(x,y) = \begin{cases} 0.000002\, e^{-0.001x - 0.002y} & 0 < x,\ 0 < y \\ 0 & \text{otherwise} \end{cases}$$

Reminder: Intersection

- In general, for two events, *E* and *F*, we have

$$P(EF) = P(E \mid F) \cdot P(F).$$

- If the events are, however, *independent*, then probability of the intersection is

$$P(EF) = P(E) \cdot P(F).$$

Independent Random Variables

- We call RV X and Y ***independent***, if for any two sets of real numbers A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}.$$

- This implies that

$$F(x, y) = F_X(x) \cdot F_Y(y), \quad x \in R, y \in R.$$

$$p(x, y) = p_X(x) \cdot p_Y(y) \text{ (discrete case)},$$

where p_X and p_Y are PMFs of X and Y.

$$f(x, y) = f_X(x) \cdot f_Y(y) \text{ (continuous case)}.$$

Food for Thought – Example 3.1

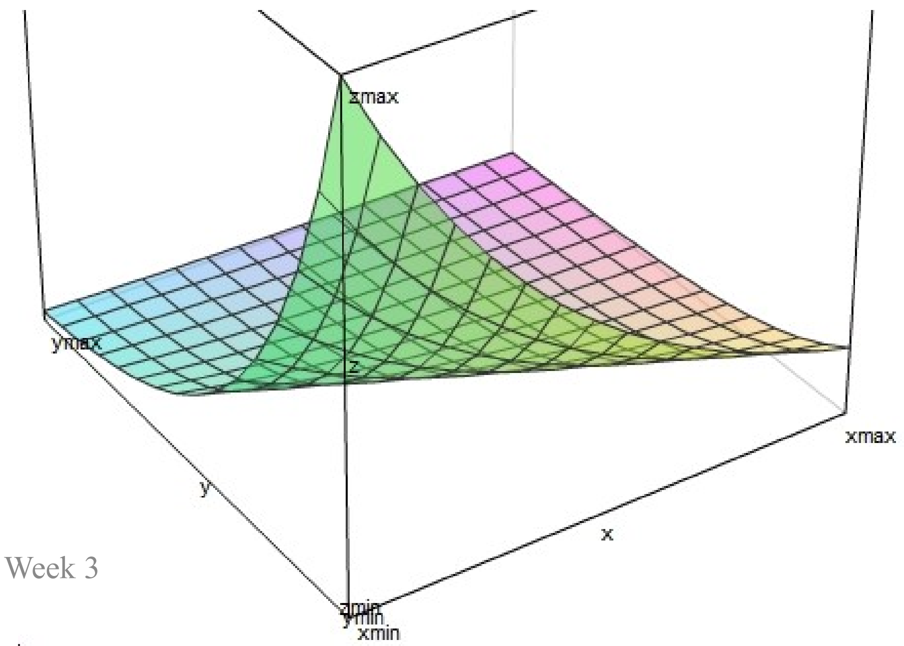
- A set of k coupons is collected. Each coupon is one of n types. Each coupon is independently a type j coupon with probability p_j such that $\sum_{j=1}^{n} p_j = 1$. Find the probability that the set contains either a type 1 or type 2 coupon.

Food for Thought – Example 3.2

- Suppose the joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x, 0 < y \\ 0 & \text{otherwise} \end{cases}$$

- Compute
  (a) P{X>1, Y<1}
  (b) P{X<Y}



Food for Thought – Example 3.3

- Decide if the following is the joint density function of independent RVs.

$$f(x,y) = \frac{1}{4\pi} e^{-\left(\frac{x^2}{2}+\frac{y^2}{8}\right)}, \qquad x, y \in \mathbb{R}$$

References/Resources Used

- Lecture Slides for MATH 401 of Dr. Oleksiy Us, Department of Mathematics, German University of Cairo. [PPT]

Food for Thought – Example 3.1

- A set of k coupons is collected. Each coupon is one of n types. Each coupon is independently a type j coupon with probability p_j such that $\sum_{j=1}^{n} p_j = 1$. Find the probability that the set contains either a type 1 or type 2 coupon.

Solution

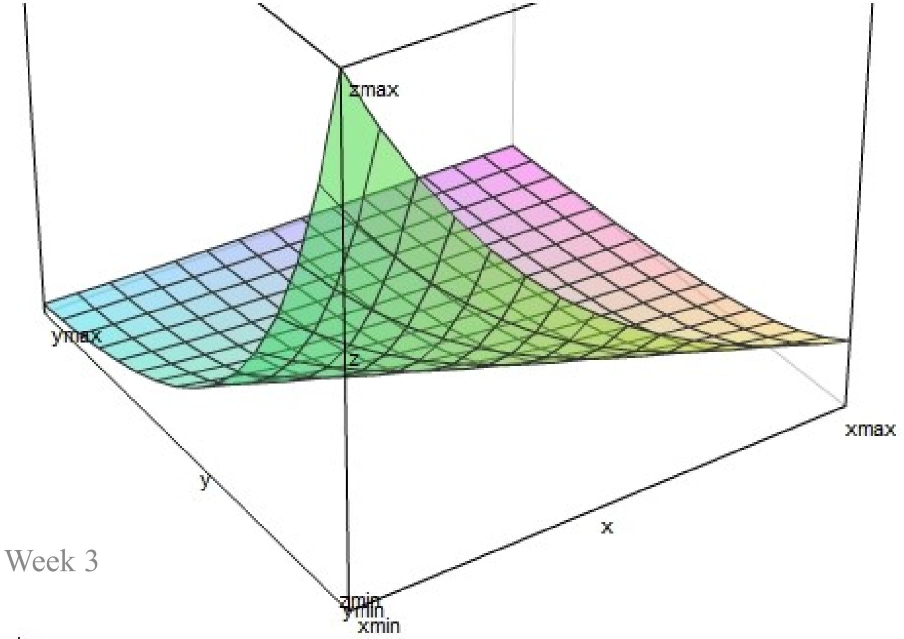
- C_i : i^{th} coupon in the set is of type 1 or 2
- $C = \bigcup_{i}^{k} C_i$
- Probability C_i is neither type 1 nor 2 is (1-p1-p2) $= P(C_i^c)$
- Probability that none of the coupons in the set is of type 1 or 2, by independence is: $\prod_{i}^{k} P(C_i^c)$ = $(1-p1-p2)^k$
- Probability that the set has at least one coupon of type 1 or 2 is therefore 1 - $(1-p1-p2)^k$

Food for Thought – Example 3.2

- Suppose the joint density function of X and Y is given by

$$f(x,y)=\begin{cases} 2e^{-x}e^{-2y} & 0<x,\ 0<y \\ 0 & \text{otherwise} \end{cases}$$

- Compute
  (a) P{X>1, Y<1}
  (b) P{X<Y}



Solution for Food for Thought - Example 3.2

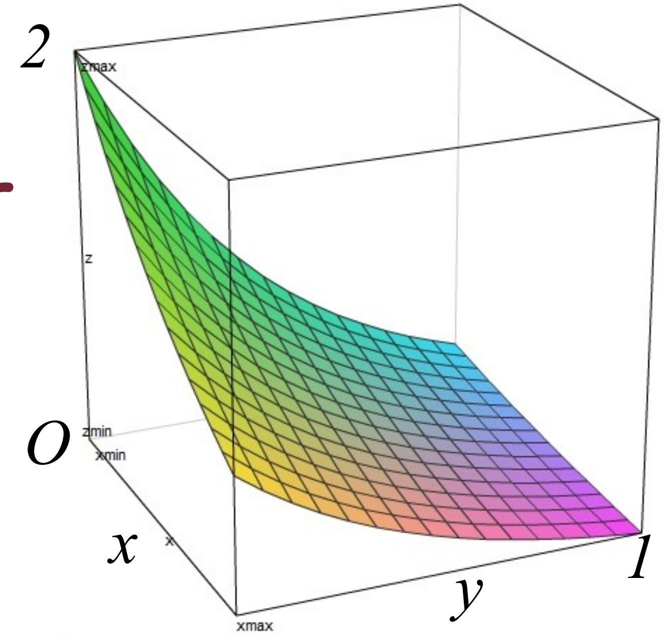
(a) $$P\{X > 1, Y < 1\} = \int_0^1 \int_1^\infty 2e^{-x} e^{-2y} \, dx dy$$

$$= \int_0^1 2e^{-2y} \left(-e^{-x} \Big|_1^\infty \right) dy$$

$$= e^{-1} \int_0^1 2e^{-2y} dy$$

$$= e^{-1}(1 - e^{-2})$$

Food for Thought



(b) $$P\{X < Y\} = \int_0^\infty \int_0^y 2e^{-x} e^{-2y} \, dx dy$$

$$= \int_0^\infty 2e^{-2y} \left(-e^{-x} \Big|_0^y \right) dy = \int_0^\infty 2e^{-2y} \left(-e^{-y} + 1 \right) dy$$

$$= \int_0^\infty 2e^{-2y} dy - \int_0^\infty 2e^{-3y} dy$$

$$= 1 - \frac{2}{3} = \frac{1}{3}$$

Food for Thought – Example 3.3

- Decide if the following is the joint density function of independent RVs.

$$f(x,y) = \frac{1}{4\pi} e^{-\left(\frac{x^2}{2}+\frac{y^2}{8}\right)}, \qquad x, y \in \mathbb{R}$$

Solution Food for Though Example 3.3

- We observe that

$$f(x,y) = \frac{1}{4\pi} e^{-\left(\frac{x^2}{2}+\frac{y^2}{8}\right)}$$

$$= \left(\frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x^2}{2}\right)}\right)\left(\frac{1}{\sqrt{8\pi}} e^{-\left(\frac{y^2}{8}\right)}\right)$$

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

Solution

- Can also confirm that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{4\pi}e^{-\left(\frac{x^2}{2}+\frac{y^2}{8}\right)}dxdy = 1$$

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{x^2}{2}\right)}dx = 1$$

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{8\pi}}e^{-\left(\frac{y^2}{8}\right)}dy = 1$$

So $f(x, y)$ is the product of two PDFs of two independent RVs.