Module 3: Random Variables

Ruben Zamar
Department of Statistics
UBC

January 18, 2016

USAGE

RANDOM VARIABLES

ARE USED TO REPRESENT

NUMERICAL FEATURES

OF A RANDOM EXPERIMENT

EXAMPLES OF RANDOM QUANTITIES

X	=	NUMBER OF DEFECTIVE ITEMS IN A LOT
Y	=	NUMBER OF VISITS TO A WEBSITE
T	=	TIME TO OCCURRENCE OF A RARE EVENT
V	=	PERCENTAGE YIELD OF A CHEMICAL PROCESS
Z	=	VERTICAL DISTANCE TO A TARGET

DEFINITION AND NOTATION

A RANDOM VARIABLE IS A FUNCTION DEFINED ON THE SAMPLE SPACE:

$$X : \Omega \to R$$

$$X(\omega) = x$$

RANDOM VARIABLES ARE DENOTED BY

$$X, \quad Y, \quad Z, \quad U, \quad V$$

LAST UPPER CASE LETTERS
IN THE ALPHABET

EXAMPLE

- **EXPERIMENT**: FLIPPING A COIN 10 TIMES
- **SAMPLE SPACE** Ω : ALL POSSIBLE SEQUENCE OF TEN HEADS (H) AND TAILS (T)
- **RANDOM VARIABLE** X : NUMBER OF HEADS
- **RANDOM VARIABLE** Y : LARGEST RUN OF TAILS
- SUPPOSE

$$\begin{aligned} \omega &= (HTTHHTTTHT) \\ X(\omega) &= 4 \\ Y(\omega) &= 3 \end{aligned}$$

MORE NOTATION

POSSIBLE VALUES OF

$$X, Y, Z, U, V,$$

ARE DENOTED BY

$$x, \quad y, \quad z, \quad u, \quad v,$$

(CORRESPONDING LOWER-CASE LETTERS)

DISCUSSION

X $=$ RANDOM QUANTITY

AND

x $=$ VALUE OF X (ω) ,

KNOWN AFTER **THE EXPERIMENT IS**

PERFORMED AND ω IS DETERMINED

DISCUSSION

THE EVENT

"RANDOM VARIABLE $\mathbf{X}$ **TAKES THE VALUE $\mathbf{x}$".**

IS MATHEMATICALLY REPRESENTED AS

$$\mathbf{X} = \mathbf{x}$$

MORE PRECISELY

$$\mathbf{X} = \mathbf{x} \quad \text{MEANS} \quad \{\omega : \mathbf{X}(\omega) = \mathbf{x}\}$$

RANGE OF A RANDOM VARIABLE

SET OF ALL THE POSSIBLE VALUES
THAT THE RANDOM VARIABLE CAN TAKE ON

		RANGE
X =	**NUMBER OF DEFECTIVE ITEMS IN A LOT**	$\{0, 1, 2, ..., N\}$
Y =	**NUMBER OF VISITS TO A WEBSITE**	$\{0, 1, 2, ...\}$
T =	**TIME TO OCCURRENCE OF A RARE EVENT**	$[0, \infty)$
V =	**PERCENTAGE YIELD OF A CHEMICAL PROCESS**	$[0, 100]$
Z =	**VERTICAL DEVIATION FROM A TARGET**	$(-\infty, \infty)$

DISCRETE RANDOM VARIABLES

A RANDOM VARIABLE IS DISCRETE WHEN ITS RANGE IS EITHER

FINITE (e.g. $\{0, 1, 2, ..., 100\}$)

OR

COUNTABLE (e.g. $\{1, 2, 3, ...\}$)

A RANDOM VARIABLE IS CONTINUOUS WHEN ITS RANGE IS AN INTERVAL

INTERVAL OF FINITE LENGTH
SUCH AS: $(0, 1)$, $[1, 5)$, $[0, 100]$

OR

INTERVAL OF INFINITE LENGTH
SUCH AS: $(0, \infty)$, $[0, \infty)$, $(-\infty, \infty)$

DISCRETE RANDOM VARIABLES

PROBABILITY MASS FUNCTION (pmf)

THE **pmf** OF A **DISCRETE** RANDOM VARIABLE X GIVES THE **PROBABILITY** OF OCCURRENCE FOR EACH POSSIBLE VALUE x OF X.

IN MATHEMATICAL SYMBOLS:

$$f(x) = P(X = x)$$

PROPERTIES OF THE pmf

1. $0 \leq f(x) \leq 1$

2. $\sum f(x) = 1$

3. $P(X \in A) = \sum_{x \in A} f(x)$

DISTRIBUTION FUNCTION (cdf)

DISTRIBUTION FUNCTIONS ARE DENOTED BY UPPER CASE LETTERS SUCH AS

F, G , H

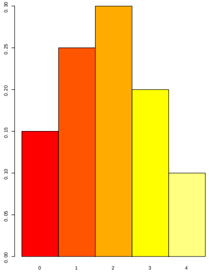
$$F(x) = P(X \leq x) = \sum_{k \leq x} f(k)$$

PROPERTIES OF THE cdf

1. $0 \leq F(x) \leq 1$
2. $F(x)$ is non decreasing
3. $F(-\infty) = 0, \quad F(\infty) = 1$
4. $P(a < X \leq b) = F(b) - F(a)$
5. $f(k) = F(k) - F(k-1)$

GIVEN BY A TABLE:

x	$f(x)$	$F(x)$
0	0.15	0.15
1	0.25	0.40
2	0.30	0.70
3	0.20	0.90
4	0.10	1.00



EXAMPLE 1 (continued)

x	$f(x)$	$F(x)$
0	0.15	0.15
1	0.25	0.40
2	0.30	0.70
3	0.20	0.90
4	0.10	1.00

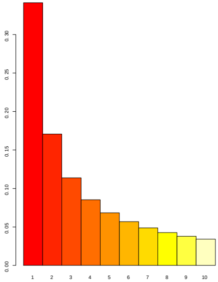
- $P(1 < X \leq 3) = F(3) - F(1) = 0.90 - 0.40 = 0.50$

- $P(1 \leq X < 3) = P(0 < X \leq 2) = F(2) - F(0) = 070 - 0.15 = 0.55$

EXAMPLE 2

GIVEN BY A FORMULA

$$f(x) = \frac{1}{2.928968} \times \frac{1}{x}, \qquad x = 1, 2, ..., 10$$



EXAMPLE 2

$$F(x) = \frac{1}{2.928968}\sum_{k=1}^{x}\frac{1}{k}, \qquad x = 1, 2, ..., 10$$

$$\begin{aligned} P(2 < X \leq 5) &= \frac{1}{2.928968}\sum_{k=1}^{5}\frac{1}{k} - \frac{1}{2.928968}\sum_{k=1}^{2}\frac{1}{k} \\ &= 0.267\,44 \end{aligned}$$

EXAMPLE 3

GIVEN BY A TABLE OF GENERIC VALUES:

x	$f(x)$	$F(x)$
0	$(1-p)^2$	$(1-p)^2$
1	$2p(1-p)$	$1-p^2$
2	p^2	1

, $0 < p < 1$

$$f(x) = \begin{cases} (1-p)^2 & x = 0 \\ 2p(1-p) & x = 1 \\ p^2 & x = 2 \end{cases}$$

- The **parameter** p can be chosen to obtain a desired configuration of probabilities.

EXPECTED VALUE

- CONSIDER A FUNCTION $g(X)$

$$
\begin{aligned}
g(X) &= X \\
g(X) &= (X-t)^2, \quad \text{for some constant value } t \\
g(X) &= e^{tX}, \quad \text{for some constant value } t
\end{aligned}
$$

- THE OPERATOR "EXPECTED VALUE" (DENOTED BY E) IS DEFINED AS FOLLOWS

$$E[g(X)] = \sum_x g(x) \; f(x)$$

x	$f(x)$
0	0.15
1	0.25
2	0.30
3	0.20
4	0.10

Take $g(X) = X^2$

$$
\begin{aligned}
E\left[X^2\right] &= \sum_{x=0}^{4} x^2 f(x) \\
&= 0 \times 0.15 + 1 \times 0.25 + 4 \times 0.30 + 9 \times 0.20 + 16 \times 0.10 = 4.85
\end{aligned}
$$

DISCUSSION

- $E(g(X))$ IS THE WEIGHTED AVERAGE OF THE FUNCTION $g(X)$
- MORE LIKELY VALUES OF $g(x)$ (WITH LARGER $f(x)$) HAVE MORE WEIGHT
- $E(g(X))$ IS CONSIDERED A "TYPICAL VALUE" OF $g(X)$, WHICH CAN BE USED TO SUMMARIZE $g(X)$.

LAW OF LARGE NUMBERS

- Suppose that $X_1, X_2, X_3, ..., X_n$ are independent measurements of the random variable X.

- Example: $X =$ number of traffic accidents in Vancouver in one week, and $X_1, X_2, X_3, ..., X_n$ are the number of traffic accidents in n consecutive weeks.

- Then, it can be shown that, as $n \to \infty$

$$\overline{X} = \frac{1}{n}\left(X_1 + X_2 + X_3 + \cdots + X_n\right) \to E\left(X\right)$$

- THE OPERATOR E IS A "LINEAR OPERATOR"

$$\begin{aligned} E\left[a+bg\left(X\right)\right] &= \sum_x \left(a+bg\left(X\right)\right)\ f\left(x\right) \\ &= \sum_x a\ f\left(x\right)+\sum_x bg\left(X\right)\ f\left(x\right) \\ &= a\overbrace{\sum_x\ f\left(x\right)}^{1}+b\overbrace{\sum_x g\left(X\right)\ f\left(x\right)}^{E[g(X)]} \\ &= a+bE\left[g\left(X\right)\right] \end{aligned}$$

MOMENTS

$g(X) = X^k, \;\; k = 1, 2, 3, ..$

$$\mu_k \;\; = \;\; E\left(X^k\right) = \sum_x x^k \; f(x)$$

Moment generating function

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} \; f(x)$$

USAGE

$$\frac{d}{dt} M_X(t)\,|_{t=0} = M_X'(0) = \mu_1$$

$$\frac{d^2}{dt^2} M_X(t)\,|_{t=0} = M_X''(0) = \mu_2$$

In general:

$$\frac{d^k}{dt^k} M_X(t)\,|_{t=0} = M_X^{(k)}(0) = \mu_k$$

MEAN, VARIANCE, STANDARD DEVIATION

KEY SUMMARY FEATURES FOR A RANDOM VARIABLE X ARE:

THE **MEAN**

$$\mu = E(X) = \sum x f(x)$$

THE **VARIANCE**

$$\sigma^2 = Var(X) = E\left[(X-\mu)^2\right] = \sum (x-\mu)^2 f(x)$$

THE **STANDARD DEVIATION**

$$\sigma = SD(X) = \sqrt{\sum (x-\mu)^2 f(x)}$$

VARIANCE FORMULA

$$
\begin{aligned}
\sigma^2 &= Var(X) \\
&= E(X^2) - E(X)^2 \\
&= \mu_2 - \mu^2
\end{aligned}
$$

VARIANCE FORMULA

PROOF:

$$\begin{aligned}
\sigma^2 &= Var(X) = E\left[(X-\mu)^2\right] = \sum (x-\mu)^2 f(x) \\
&= \sum \left(x^2 + \mu^2 - 2\mu x\right) f(x) \\
&= \sum x^2 f(x) + \sum \mu^2 f(x) - \sum 2\mu x f(x) \\
&= \overbrace{\sum x^2 f(x)}^{E(X^2)} + \mu^2 \overbrace{\sum f(x)}^{=1} - 2\mu \overbrace{\sum x f(x)}^{=\mu} \\
&= E\left(X^2\right) + \mu^2 - 2\mu^2 = \mu_2 - \mu^2
\end{aligned}$$

EXAMPLE 1 (continued)

x	$f(x)$	$xf(x)$	$x^2 f(x)$
0	0.15	0.00	0.00
1	0.25	0.25	0.25
2	0.30	0.60	1.20
3	0.20	0.60	1.80
4	0.10	0.40	1.60
Total	—	$\mu = 1.85$	$\mu_2 = 4.85$

$$\mu = 1.85$$

$$\sigma^2 = \mu_2 - \mu^2 = 4.85 - 1.85^2 = 1.4275$$

$$\sigma = \sqrt{1.4275} = 1.1948$$

EXAMPLE 2 (continued)

$$f(x) = \frac{1}{2.928968} \times \frac{1}{x}, \qquad x = 1, 2, ..., 10$$

$$\mu = \frac{1}{2.928968} \sum_{X=1}^{10} x \frac{1}{x} = \frac{10}{2.928968} = 3.4142$$

$$\mu_2 = \frac{1}{2.928968} \sum_{X=1}^{10} x^2 \frac{1}{x} = \frac{1}{2.928968} \sum_{X=1}^{10} x$$

$$= \frac{10 \times 11}{2 \times 2.928968} = 18.77855$$

EXAMPLE 2 (continued)

$$\mu = 3.4142 \qquad \mu_2 = 18.77855$$

HENCE:

$$\sigma^2 = \mu_2 - \mu^2 = 18.77855 - 3.4142^2 = 7.1218$$

$$\sigma = \sqrt{7.1218} = 2.6687$$

EXAMPLE 3 (continued)

$$f(x) = \begin{cases} (1-p)^2 & x = 0 \\ 2p(1-p) & x = 1 \\ p^2 & x = 2 \end{cases}$$

$$\begin{aligned} \mu &= 0 \times f(0) + 1 \times f(1) + 2 \times f(2) \\ &= 2p(1-p) + 2p^2 = 2p \end{aligned}$$

EXAMPLE 3 (continued)

$$\mu_2 = 0 \times f(0) + 1 \times f(1) + 4 \times f(2)$$
$$= 2p(1-p) + 4p^2 = 2p + 2p^2$$

$$\sigma^2 = \mu_2 - \mu_1^2 = \left(2p + 2p^2\right) - (2p)^2$$
$$= 2p - 2p^2 = 2p(1-p)$$

$$\sigma = \sqrt{2p(1-p)}$$

EXAMPLE 3 (continued)

p	μ	σ^2	σ
0.05	0.10	0.095	0.308
0.25	0.50	0.375	0.612
0.50	1.00	0.500	0.707
0.75	1.50	0.375	0.612
0.95	1.90	0.095	0.308

EXAMPLE 4: An urn contains n chips numbered 1 through n. We draw k chips $(1 < k < n)$ without replacement. Let Y represent the highest number among those drawn.

(a) What is the range of Y?

(b) Find $F_Y(y)$.

(c) Find $f_Y(y)$.

(d) Suppose $n = 20$ and $k = 5$. Calculate the mean, variance and standard deviation for Y.

EXAMPLE 4 (Cont)

(a) **What is the range of Y?**

Smallest possible value of Y is k.

[corresponds to the event $\{1, 2, ..., k\}$]

Largest possible value of Y is n

$$Range = \{k, k+1, ..., n\}$$

EXAMPLE 4 (Cont)

(b) **Find** $F(y)$

$$F(y) = \frac{\overbrace{\binom{y}{k}}^{\text{Because the } k \text{ chips must be } \leq y}}{\binom{n}{k}} = \frac{\frac{y!}{k!(y-k)!}}{\frac{n!}{k!(n-k)!}}$$

$$= \frac{(n-k)!y!}{n!(y-k)!}, \quad y = k, k+1, ..., n$$

(c) **Find** $f(y)$.

$$f(y) = F(y) - F(y-1), \quad y = k, k+1, ..., n$$

NOTE: $F(y) = 0$ for all $y < k$. So

$$f(k) = F(k) - F(k-1) = F(k)$$

$$= \frac{\binom{k}{k}}{\binom{n}{k}} = \frac{1}{\binom{n}{k}} = \frac{k!(n-k)!}{n!}$$

EXAMPLE 4 (Cont)

(d) Suppose $n = 20$ and $k = 5$. Calculate the mean, variance and standard deviation for Y.

This must be done using a computer (e.g. R or matlab)

PROPERTIES OF THE MEAN

Let

$$\mu_X = E(X), \quad \mu_Y = E(Y)$$

- $E(a + bX) = a + bE(X) = a + b\mu_X,$ for constants a, b.

- $E(a + bX + cY) = a + b\mu_X + c\mu_Y,$ for constants a, b, c.

PROPERTIES OF THE MEAN

The mean minimizes the Mean Square Error:

$$S(t) = E\left[(X-t)^2\right] \geq E\left[(X-\mu)^2\right] = Var(X), \quad \text{for all } t$$

Proof:

$$\begin{aligned} S(t) &= E\left(X^2 + t^2 - 2Xt\right) = E\left(X^2\right) + t^2 - 2\mu t \\ S'(t) &= 2t - 2\mu = 0 \Rightarrow t = \mu \\ S''(\mu) &= 2 > 0 \ \ (\mu \text{ is a minimizer}) \end{aligned}$$

PROPERTIES OF THE VARIANCE

$$
\begin{aligned}
Var(a + bX) &= E\left[a + bX - E(a + bX)\right]^2 \\
&= E\left[a + bX - a - E(bX)\right]^2 = E\left[bX - bE(X)\right]^2 \\
&= E\left[b^2 (X - E(X))^2\right] = b^2 E\left[(X - E(X))^2\right] \\
&= b^2 Var(X)
\end{aligned}
$$

That is:

$$
\begin{aligned}
Var(a + bX) &= b^2 Var(X) \\
SD(a + bX) &= |b| \, SD(X)
\end{aligned}
$$

BINOMIAL RANDOM VARIABLES

- Suppose we wish to monitor the occurrence of an event A, such as

$$A = \{\text{Electric circuit has a flaw}\}$$

$$A = \{\text{Wind speed exceeds 100k}\}$$

$$A = \{\text{Student passes a given math test}\}$$

- Let

$$p = P(A)$$

- The occurrence of A is arbitrarily called a "success"
- The non-occurrence is arbitrarily called a "failure"

BINOMIAL RANDOM VARIABLES (continued)

- The occurrence/non-occurrence of A is monitored a fixed number, n, of times
- Each monitoring is called "a trial".
- **INDEPENDENCE:** We perform n independent trials
- The random quantity of interest:

$$X = \text{Number of successes}$$

- Notation: $X \sim Bin\,(n, p)$,
  $n =$ number of trials
  $p =$probability of success.

BINOMIAL RANDOM VARIABLES (continued)

- Possible values for X are:

$$\text{Range} = \{0, 1, 2, ..., n\}$$

- The Binomial density:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, ..., n$$

- The combinatorial coefficient:

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

MEAN AND VARIANCE

- Moment generating function

$$M(t) = \left(1 - p + pe^t\right)^n$$

- Mean

$$\mu = E(X) = np$$

- Variance

$$\sigma^2 = np(1-p), \quad \text{maximized when} \quad p = 1/2$$

MEAN AND VARIANCE (Continued)

$$M'(t) = \frac{d}{dt}M(t) = \frac{d}{dt}\left(1-p+pe^t\right)^n = n\left(1-p+pe^t\right)^{n-1}pe^t$$

$$M''(t) = n(n-1)\left(1-p+pe^t\right)^{n-1}p^2e^{2t} + n\left(1-p+pe^t\right)^{n-1}pe^t$$

Therefore,

$$\mu = M'(0) = n\left(1-p+pe^0\right)^{n-1}pe^0 = np$$

$$\mu_2 = n(n-1)p^2 + np$$

$$\sigma^2 = \mu_2 - \mu^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$$

PRACTICE

Problem: suppose that finding oil when digging at certain locations has probability $p = 0.10$ (geologically determined locations).

(a) How many wells should we dig to find oil with probability larger than or equal to 0.95?

(b) How many wells should we dig to obtain at least 2 successful wells with probability larger than or equal to 0.95?

PRACTICE (continued)

Solution Part (a)

Assume the diggings are independent. Hence the number of successful wells is $X \sim Bin(n, 0.10)$, where n is the number of dug wells.

$$\begin{aligned} P(X > 0) &= 1 - P(X = 0) = 1 - (1 - 0.10)^n = 0.95 \\ (1 - 0.10)^n &= 1 - 0.95 \\ n \ln(0.90) &= \ln(0.05) \Longrightarrow n = \frac{\ln(0.05)}{\ln(0.90)} = 28.43, \end{aligned}$$

Answer: $n = 29$.

PRACTICE (continued)

Solution Part (b)

Assume the diggings are independent. Hence the number of successful wells is $X \sim Bin(n, 0.10)$, where n is the number of dug wells.

$$P(X > 1) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - (1 - 0.10)^n - n0.10(1 - 0.10)^{n-1} \geq 0.95$$

$$0.9^n + n \times 0.1 \times 0.9^{n-1} \leq 0.05$$

This equation must be solved numerically. **Answer:** $n = 46$

Suppose we wish to count the number of occurrences of a certain event A, such as

$$A = \{\text{Earthquakes over 5.0 in BC in one year}\}$$

$$A = \{\text{Traffic violations at Oak \& Cambie in one week}\}$$

$$A = \{\text{Rainfalls exceeding 30mm in Vancouver in one year}\}$$

THE RATE OF OCCURRENCE OF A

Let λ be the rate of occurrence for the event of interest, such as

$$\lambda = 4 \text{ per year}$$

$$\lambda = 15 \text{ per week}$$

- Number of occurrences: 4, 15, etc.
- The time interval for the count MUST BE TAKEN INTO ACCOUNT: year, week, etc.

ASSUMPTIONS

NOTATION: $P(k;t)$ is the probability of k occurrences of A in the interval $[0,t]$

1. **INDEPENDENCE:** occurrences in **disjoint** time intervals are **independent**

2. **PROPORTIONALITY:**

$$P(1;t) = \lambda t + o(t), \quad \text{where} \quad \lim_{t \to 0} \frac{o(t)}{t} = 0$$

3. **RARE EVENT:** We have at most 1 occurrence of A in a small period of time

$$1 - P(0;t) - P(1;t) = \sum_{k=2}^{\infty} P(k;t) = o(t)$$

POISSON PROBABILITY MASS FUNCTION (pmf)

The quantity of interest is:

$$X = \text{Number of occurrences}$$

The possible values for X are:

$$\text{Range} = \{0,1,2,...\}$$

The Poisson **pmf** is:

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, ...$$

POISSON MEAN AND VARIANCE

- Moment Generating Function

$$M(t) = e^{\lambda(e^t - 1)}$$

- Mean

$$\mu = E(X) = \lambda$$

- Variance

$$\sigma^2 = Var(X) = \lambda$$

COMPUTING THE MEAN

$$\mu = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda \overbrace{\sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!}}^{=1} = \lambda$$

THE MOMENT GENERATING FUNCTION

$$
\begin{aligned}
M(t) &= \sum_{x=0}^{\infty} e^{-tx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{\lambda e^t} e^{-\lambda} \sum_{x=0}^{\infty} e^{-\lambda e^t} \frac{(\lambda e^t)^x}{x!} = e^{\lambda(e^t-1)} \overbrace{\sum_{x=0}^{\infty} \frac{e^{-\lambda e^t} (\lambda e^t)^x}{x!}}^{=1} \\
&= e^{\lambda(e^t-1)}
\end{aligned}
$$

PRACTICE

Problem: Suppose that the number Y of earthquakes over 5.0 (Richter scale) in a given area is a Poisson random variable $[Y \sim \mathcal{P}(\lambda)]$ with $\lambda = 3.6$ per year.

1. What is the probability of having at least 2 earthquakes over 5.0 during the next 6 months?
2. What is the probability of having 1 earthquake over 5.0 next month?
3. What is the probability of waiting more than 3 months for the next earthquake over 5.0 in that area?

Solution:

We should keep track of the length of the period of interest to adjust the rate:

$$\boxed{3.6 \text{ per year} = 1.8 \text{ per half year} = 0.3 \text{ per month}}$$

PRACTICE

Solution Part 1:

$$X = \# \text{ of earthquakes in the next 6 month} \sim \mathcal{P}(1.8)$$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \frac{e^{-1.8} \times 1.8^0}{0!} - \frac{e^{-1.8} \times 1.8}{1!} \\ &= 1 - e^{-1.8} - e^{-1.8} \times 1.8 = 0.53716 \end{aligned}$$

PRACTICE

Solution Part 2:

$$X = \# \text{ of earthquakes in the next month} \sim \mathcal{P}(0.3)$$

$$P(X = 1) = e^{-0.3} \times 0.3 = 0.22225$$

Solution Part 3:

$$X = \# \text{ of earthquakes in the next quarter} \sim \mathcal{P}(0.9)$$

$$P(\text{Waiting more than 3 months}) = P(X = 0) = e^{-0.9} = 0.40657$$

CONTINUOUS RANDOM VARIABLES

CONTINUOUS DENSITY

1. $f(x) \geq 0,$ **NON-NEGATIVE**

2. $\int_{-\infty}^{\infty} f(x)\ dx = 1,$ **INTEGRATES TO ONE**

3. $P(a < X < b) = \int_{a}^{b} f(x)\ dx,$ **USED TO COMPUTE PROBABILITIES**

CONTINUOUS DISTRIBUTION FUNCTION

CONTINUOUS DISTRIBUTION FUNCTION

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$$

F CAN BE USED TO COMPUTE $P(a < X < b)$:

$$\begin{aligned} P(a < X < b) &= \int_{a}^{b} f(x)\, dx = \int_{-\infty}^{b} f(x)\, dx - \int_{\infty}^{a} f(x)\, dx \\ &= F(b) - F(a) \end{aligned}$$

DISCUSSION

NOTE 1: IN THE CONTINUOUS CASE

$$P(X = x) \quad = \quad \int_x^x f(t)\,dt = 0 \qquad \text{FOR ALL} \quad x$$

NOTE 2:

$$f(x) \quad \neq \quad P(X = x)$$

IN PARTICULAR, WE OFTEN HAVE

$$f(x) > 1$$

DISCUSSION

NOTE 3: FOR SMALL $\delta > 0$,

$$P(x < X < x + \delta) = \int_x^{x+\delta} f(t)\, dt$$

$$\approx f(x)\delta$$

NOTE 4:

$$F'(x) = \frac{d}{dx} \int_{-\infty}^{x} f(t)\, dt = f(x)$$

DISCUSSION

NOTE 5: SINCE

$$P(X = a) = P(X = b) = 0,$$

WE HAVE

$$\begin{aligned} P(a < X < b) &= P(a < X \leq b) = P(a \leq X < b) \\ &= P(a \leq X \leq b) = F(b) - F(a) \end{aligned}$$

MEAN, VARIANCE AND STANDARD DEVIATION

MEAN:

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

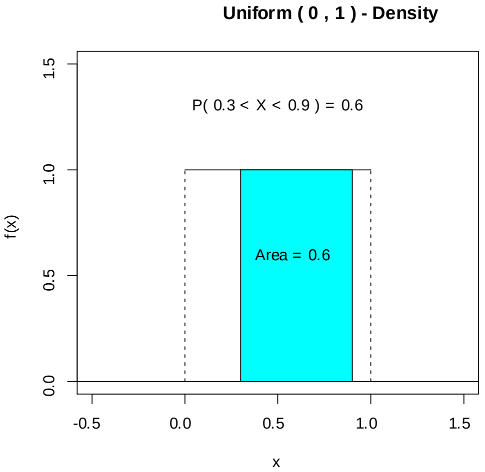
VARIANCE:

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \, f(x) \, dx = \int_{-\infty}^{\infty} x^2 \, f(x) \, dx - \mu^2 = \mu_2 - \mu^2$$

STANDARD DEVIATION

$$\sigma = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 \, f(x) \, dx}$$

UNIFORMLY DISTRIBUTED RANDOM VARIABLES



UNIFORM DENSITY (continued)

- Notation: $X \sim Unif\,(\alpha, \beta)$
- Parameters: α (lower limit) and β (upper limit).
- Naturally, $\alpha < \beta$.
- In the picture (Example 1) we have

$$\alpha = 0 \text{ and } \beta = 1$$

UNIFORM DENSITY (continued)

- Mathematical representation of the density:

$$f(x) = \begin{cases} 0 & x \leq \alpha \\ 1/(\beta - \alpha) & \alpha < x < \beta \\ 0 & x \geq \beta \end{cases}$$

- We will often use the shorter notation

$$f(x) = 1/(\beta - \alpha), \quad \alpha < x < \beta$$

UNIFORM DISTRIBUTION FUNCTION

$$F(x) = \begin{cases} 0 & x \leq \alpha \\ \frac{1}{\beta-\alpha}\int_{\alpha}^{x} dt = \frac{x-\alpha}{\beta-\alpha} & \alpha < x < \beta \\ 1 & x \geq \beta \end{cases}$$

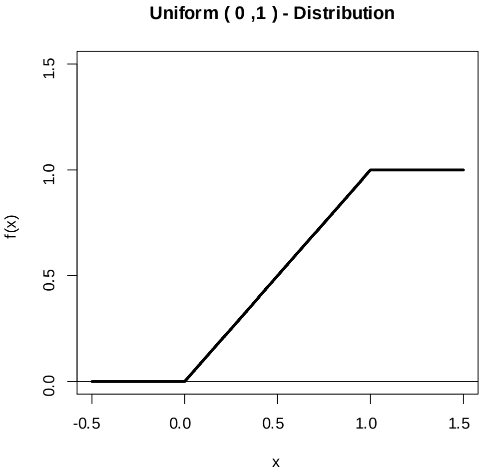
Shorter notation:

$$F(x) = \frac{x-\alpha}{\beta-\alpha}, \qquad \alpha < x < \beta$$

Hence,

$$P(a < X < b) = F(b) - F(a) = \frac{b-a}{\beta-\alpha}$$

UNIF(0,1) DISTRIBUTION FUNCTION



UNIFORM DENSITY - SUMMARY MEASURES

First and second moments:

$$\mu = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} x dx = \frac{\alpha + \beta}{2}$$

$$\mu_2 = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} x^2 dx = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha^2 + \alpha\beta}{3}$$

Variance:

$$\mu = \frac{\alpha+\beta}{2} \quad , \quad \mu_2 = \frac{\beta^2+\alpha^2+\alpha\beta}{3}$$

$$\begin{aligned}\sigma^2 &= \mu_2 - \mu^2 = \frac{b^2+\alpha^2+\alpha\beta}{3} - \frac{b^2+\alpha^2+2\alpha\beta}{4} \\ &= \frac{4\left(b^2+\alpha^2+\alpha\beta\right) - 3\left(b^2+\alpha^2+2\alpha\beta\right)}{12} \\ &= \frac{b^2+\alpha^2-2\alpha\beta}{12} = \frac{(\beta-\alpha)^2}{12}\end{aligned}$$

PRACTICE

Problem: Suppose that X is $\text{Unif}(0, 10)$. Calculate $P(X > 3)$ and $P(X > 5 | X > 2)$.

Solution:

$$
\begin{aligned}
P(X > 3) &= 1 - F(3) \\
&= 1 - \frac{3}{10} = 0.70
\end{aligned}
$$

PRACTICE

$$
\begin{aligned}
P\left(X>5 \mid X>2\right) &= \frac{P\left(\{X>5\} \cap \{X>2\}\right)}{P\left(X>2\right)} \\
&= \frac{P\left(X>5\right)}{P\left(X>2\right)} = \frac{1-F\left(5\right)}{1-F\left(2\right)} \\
&= \frac{1-\left(5/10\right)}{1-\left(2/10\right)} = 0.625
\end{aligned}
$$

PRACTICE

Problem (Change of Variable): Suppose that X is $\text{Unif}(0,1)$. That is,

$$f_X(x) = 1 \quad 0 \le x \le 1$$

and

$$F_X(x) = x, \quad 0 \le x \le 1$$

Derive the distribution function and density function for

$$Y = -\ln(X)$$

PRACTICE

Solution: first notice that the range of Y is $(0,\infty)$, so

$$F_Y(y) = 0, \quad \text{for} \quad y < 0$$

On the other hand, for $y > 0$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(-\ln(X) \leq y) = P\left(X \geq e^{-y}\right) \\ &= 1 - F_X\left(e^{-y}\right) = 1 - e^{-y} \end{aligned}$$

PRACTICE

Hence, for $y > 0$,

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

Finally, for $y > 0$,

$$\begin{aligned} f_Y(y) &= F_Y'(y) = \frac{d}{dy}\left(1 - e^{-y}\right) \\ &= e^{-y} \end{aligned}$$

EXPONENTIAL RANDOM VARIABLES

This type of random variables are used to represent (model) the **waiting time** until the occurrence of a certain event, such as

- arrival of a customer

- occurrence of an earthquake

- crash of computer network

RATE OF OCCURRENCE

The exponential density function has a single parameter, $\lambda > 0$, which represents the **rate of occurrence** for the event

Examples:

$$\lambda = 5 \text{ per hour}$$

$$\lambda = 2 \text{ per year}$$

$$\lambda = 1 \text{ per month}$$

EXPONENTIAL DENSITY AND DISTRIBUTION

- Notation $X \sim Exp(\lambda)$
- Density function

$$f(x) = \lambda e^{-\lambda x},$$

for $x > 0$, and zero otherwise.

- Distribution function

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$$

for $x > 0$, and zero otherwise.

DISCUSSION

- Comparing with the result in the Practice above, we have

$$-\ln\left(U\left(0,1\right)\right) = Exp\left(1\right)$$

More generally,

$$-\ln\left(U\left(0,1\right)\right)/\lambda = Exp\left(\lambda\right)$$

Problem: Suppose that $X \sim Exp(\lambda)$. For $x_0 > 0$ and $h > 0$, calculate $P(X > h)$ and $P(X > x_0 + h \mid X > x_0)$. Comment on the result.

Solution:

$$P(X > h) = 1 - F(h) = e^{-\lambda h}$$

$$\begin{aligned} P(X > x_0 + h \mid X > x_0) &= \frac{P(\{X > x_0 + h\} \cap \{X > x_0\})}{P(X > x_0)} \\ &= \frac{P(X > x_0 + h)}{P(X > x_0)} = \frac{e^{-\lambda(x_0+h)}}{e^{-\lambda x_0}} \\ &= e^{-\lambda h} \end{aligned}$$

DISCUSSION

- The probability of surviving h additional units at age x is the same for all x.

- If X represents "time to failure for a system" than the system doesn't get "old".

FAILURE RATE

The failure rate is defined as

$$
\begin{aligned}
\lambda(x) &= \lim_{\delta \to 0} \frac{P(X \le x + \delta \mid X > x)}{\delta} \\
&= \lim_{\delta \to 0} \frac{P(x < X \le x + \delta)}{P(\ X > x)\,\delta} \\
&= \frac{1}{P(\ X > x)} \lim_{\delta \to 0} \frac{F(x+\delta) - F(x)}{\delta} \\
&= \frac{f(x)}{1 - F(x)} = -\frac{d}{dx} \ln(1 - F(x))
\end{aligned}
$$

HEURISTIC INTERPRETATION

$$\lambda(x)\,\delta \quad \approx \quad P\left(X \leq x+\delta \mid X > x\right), \quad \text{for} \;\; \text{small } \delta$$

Why?

Set $\qquad g(x,\delta) = P\left(X \leq x+\delta \mid X > x\right)$

Then

$$g(x,\delta) \approx \overbrace{g(x,0)}^{0} + \overbrace{\frac{\partial}{\partial\delta} g(x,0)}^{\lambda(x)}\,\delta = \lambda(x)\,\delta$$

DISCUSSION

The failure rate can be:

Constant	The system doesn't improve nor wear out with time
Increasing	The system wears out with time
Decreasing	The system improves with time

FAILURE RATE AND cdf

We have the following **result**:

$$F(x) = 1 - e^{-\int_0^x \lambda(t)dt}$$

FAILURE RATE AND cdf

Proof: Recall that $-\lambda(t) = \frac{d}{dt}\ln(1-F(t))$. Hence,

$$\begin{aligned}
-\int_0^x \lambda(t)\,dt &= \int_0^x \frac{d}{dx}\left[\ln(1-F(t))\right]dt \\
&= \ln(1-F(t))\Big|_0^x \\
&= \ln(1-F(x)) - \ln(1-F(0)) \\
&= \ln(1-F(x))
\end{aligned}$$

Therefore

$$1 - e^{-\int_0^x \lambda(t)dt} = F(x)$$

CONSTANT FAILURE RATE

$$\lambda(x) = \gamma$$

Show that in this case

$$F(x) = 1 - e^{-\gamma x} \quad \text{(exponential distribution)}$$

Increasing failure rate. For example

$$\lambda(x) = x$$

Show that for this example

$$F(x) = 1 - e^{-x^2/2} \quad \text{(Weibull distribution)}$$

DECREASING FAILURE RATE

Decreasing failure rate. For example

$$\lambda(x) = \frac{1}{1+x}$$

Show that for this example

$$F(x) = \frac{x}{1+x}$$