ME 102: Lecture 5

S. Gopalakrishnan

August 11, 2016

Empirical rule of normal datasets

If a data set is approximately normal with sample mean $\bar{x}$ and sample standard deviation s, then the following statements are true.

1. Approximately 68 percent of the observations lie within
$$\bar{x} \pm s$$
2. Approximately 95 percent of the observations lie within
$$\bar{x} \pm 2s$$
3. Approximately 99.7 percent of the observations lie within
$$\bar{x} \pm 3s$$

Example Problem

The following stem and leaf plot gives the scores on a statistics exam taken by industrial engineering students.

9	0,1,4
8	3,5,5,7,8
7	2,4,4,5,7,7,8
6	0,2,3,4,6,6
5	2,5,5,6,8
4	3,6

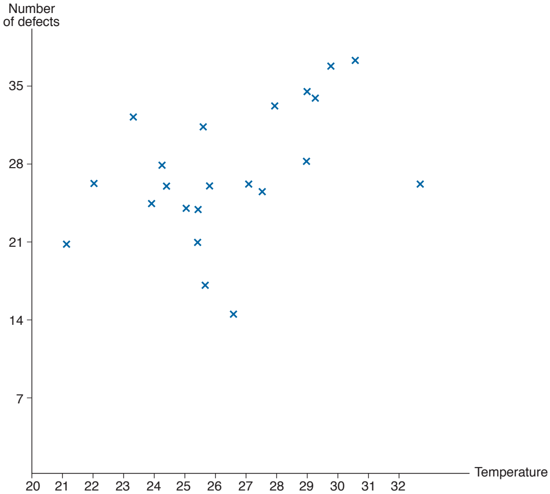
Use it to assess the empirical rule.

Paired data sets: Example Problem

TABLE 2.8 *Temperature and Defect Data*

Day	Temperature	Number of Defects
1	24.2	25
2	22.7	31
3	30.5	36
4	28.6	33
5	25.5	19
6	32.0	24
7	28.6	27
8	26.5	25
9	25.3	16
10	26.0	14
11	24.4	22
12	24.8	23
13	20.6	20
14	25.1	25
15	21.4	25
16	23.7	23
17	23.9	27
18	25.2	30
19	27.4	33
20	28.3	32
21	28.8	35
22	26.6	24

Example Problem



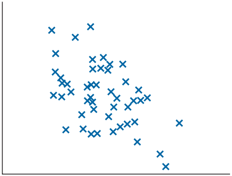
Paired data sets and correlations

Let s_x and s_y denote, respectively, the sample standard deviations of the x values and the y values. The sample correlation coefficient, call it r , of the data pairs $(x_i, y_i), i = 1, \ldots, n$ is defined by

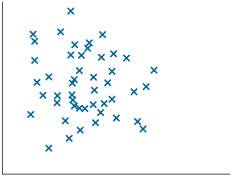
$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$
$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

When $r > 0$ we say that the sample data pairs are positively correlated, and when $r < 0$ we say that they are negatively correlated.

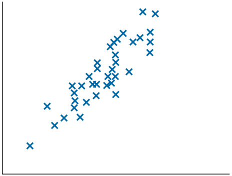
Different correlations



$r = -.50$



$r = 0$



$r = -.90$

Properties of r

1. $-1 \leq r \leq 1$
2. If for constants a and b,with $b > 0$,

$$y_i = a + bx_i, \quad i = 1, \ldots, n$$

   then $r = 1$.
3. If for constants a and b,with $b < 0$,

$$y_i = a + bx_i, \quad i = 1, \ldots, n$$

   then $r = -1$.
4. If r is sample correlation coefficient for the data pairs $x_i, y_1, i = 1, \ldots, n$ then it is also the sample correlation coefficient of the data pairs.

$$a + bx_i, \quad c + dy_i, \quad i = 1, \ldots, n$$

   provided that b and d are both positive or both negative.

Example Problem

The following data gives the resting pulse rates (in beats per minute) and the years of schooling of 10 individuals.

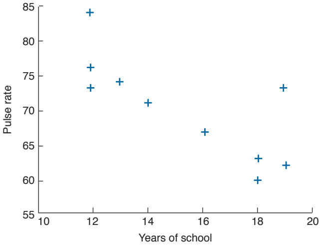
Person	1	2	3	4	5	6	7	8	9	10
Years of school	12	16	13	18	19	12	18	19	12	14
Pulse rate	73	67	74	63	7384	60	62	76	71	

Example Problem

The following data gives the resting pulse rates (in beats per minute) and the years of schooling of 10 individuals.

Person	1	2	3	4	5	6	7	8	9	10
Years of school	12	16	13	18	19	12	18	19	12	14
Pulse rate	73	67	74	63	7384	60	62	76	71	

A scatter diagram of these data is presented in the figure. The sample correlation coefficient for these data is $r = 0.7638$.



We will now prove the first three properties of the sample correlation coefficient r . That is, we will prove that $|r| \leq 1$ with equality when the data lie on a straight line. To begin, note that,

$$\sum \left(\frac{x_i - \bar{x}}{s_x} - \frac{y_i - \bar{y}}{s_y} \right)^2 \geq 0$$

$$\sum \frac{(x_i - \bar{x})^2}{s_x^2} + \sum \frac{(y_i - \bar{y})^2}{s_y^2} - 2 \sum \frac{(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y} \geq 0$$

$$n - 1 + n - 1 - 2(n-1)r \geq 0$$

showing that

$$r \leq 0$$

Note also that $r = 1$ if and only if there is equality in first equation on the previous slide. That is, $r = 1$ if and only if for all i

$$\frac{\bar{y} - y_i}{s_y} = \frac{\bar{x} - x_i}{s_x}$$

$$y_i = \bar{y} - \frac{s_y}{s_x}\bar{x} + \frac{s_y}{s_x}x_i$$

That is, $r = 1$ if and only if the data values (x_i, y_i) lie on a straight line having a positive slope.

To show that $r \geq -1$, with equality if and only if the data values (x_i, y_i) lie on a straight line having a negative slope, start with

$$\sum \left(\frac{x_i - \bar{x}}{s_x} + \frac{y_i - \bar{y}}{s_y} \right)^2 \geq 0$$

and use an argument analogous to the one just given.

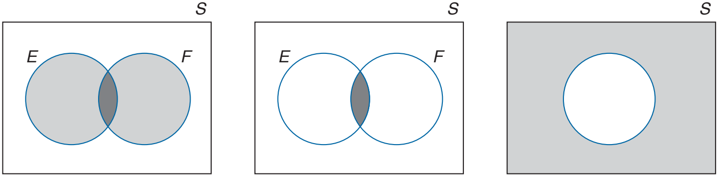
Probability

Sample space and Events

This set of all possible outcomes of an experiment is known as the **sample space** of the experiment and is denoted by S. Some examples are the following.

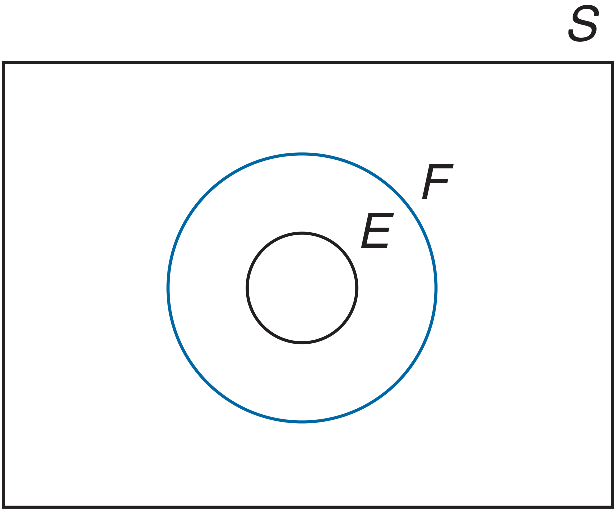
1. If the outcome of an experiment consists in the determination of the sex of a newborn child, then
$$S = g, b$$
where the outcome g means that the child is a girl and b that it is a boy.
2. If the experiment consists of the running of a race among the seven horses having post positions 1, 2, 3, 4, 5, 6, 7, then
$$S = \{\text{all orderings of } (1, 2, 3, 4, 5, 6, 7)\}$$
The outcome (2, 3, 1, 6, 5, 4, 7) means, for instance, that the number 2 horse is first, then the number 3 horse, then the number 1 horse, and so on.
3. Suppose we are interested in determining the amount of dosage that must be given to a patient until that patient reacts positively. One possible sample space for this experiment is to let S consist of all the positive numbers. That is, let
$$S = (0, \infty)$$
where the outcome would be x if the patient reacts to a dosage of value x but not to any smaller dosage.

Venn Diagrams



(a) Shaded region: $E \cup F$ (b) Shaded region: EF (c) Shaded region: E^c

Venn Diagrams



$E \subset F$

Algebra of Events

Commutative Law	$E \cup F = F \cup E$	$EF = FE$
Associative Law	$(E \cup F) \cup G = E \cup (F \cup G)$	$(EF)G = E(FG)$
Distributive Law	$(E \cup F)G = EG \cup FG$	$EF \cup G = (E \cup G)(F \cup G)$

Algebra of Events

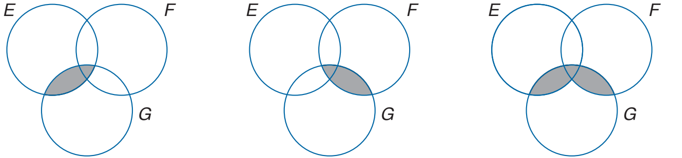
Commutative Law	$E \cup F = F \cup E$	$EF = FE$
Associative Law	$(E \cup F) \cup G = E \cup (F \cup G)$	$(EF)G = E(FG)$
Distributive Law	$(E \cup F)G = EG \cup FG$	$EF \cup G = (E \cup G)(F \cup G)$

De Morgan's Laws

$$(E \cup F)^c = E^c F^c$$

$$(EF)^c = E^c \cup F^c$$

Proof of Distributive Law



(a) Shaded region: EG (b) Shaded region: FG (c) Shaded region: $(E \cup F)G$

$(E \cup F)G = EG \cup FG$

Axioms of Probability

Axiom 1

$$0 \leq P(E) \leq 1$$

Axioms of Probability

Axiom 1

$$0 \leq P(E) \leq 1$$

Axiom 2

$$P(S) = 1$$

Axioms of Probability

Axiom 1

$$0 \leq P(E) \leq 1$$

Axiom 2

$$P(S) = 1$$

Axiom 3 For any sequence of mutually exclusive events $E_1, E_2, \ldots$ (that is,events for which $E_i E_j = \emptyset$ when $i \neq j$),

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i), \qquad n = 1, 2, \ldots, \infty$$

We call $P(E)$ the probability of the event E.

Proposition 1

Since $E \cup E^c = S$

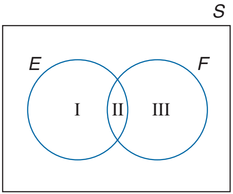
$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$$

Equivalently

$$P(E^c) = 1 - P(E)$$

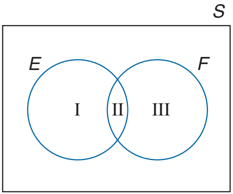
Proposition 2

$$P(E \cup F) = P(E) + P(F) - P(EF)$$



Proposition 2

$$P(E \cup F) = P(E) + P(F) - P(EF)$$



Proof

$$P(E \cup F) = P(I) + P(II) + P(III)$$
$$P(E) = P(I) + P(II)$$
$$P(F) = P(II) + P(III)$$

which shows that

$$P(E \cup F) = P(E) + P(F) - P(EF)$$

Sample spaces with equally likely outcomes

For a large number of experiments, it is natural to assume that each point in the sample space is equally likely to occur. That is, for many experiments whose sample space S is a finite set, say $S = \{1, 2, \ldots, N\}$, it is often natural to assume that

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{N\}) = p$$

Now it follows from Axioms 2 and 3 that

$$1 = P(S) = P(\{1\}) + P(\{2\}) = + \cdots + P(\{N\}) = Np$$

which shows that

$$P(\{i\}) = p = \frac{1}{N}$$

from Axiom 3 it follows that for any event E,

$$P(E) = \frac{\text{Number of points in} E}{N}$$

In words, if we assume that each outcome of an experiment is equally likely to occur, then the probability of any event E equals the proportion of points in the sample space that are contained in E.

Generalized Basic Principle of Counting

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if, . . . , then there are a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

Generalized Basic Principle of Counting

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if, . . . , then there are a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

Notation

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

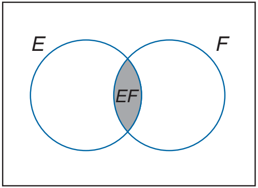
Also since $0! = 1$

$$\binom{n}{0} = \binom{n}{n} = 1$$

Conditional Probability

Conditional probability of E given that F has occurred, and is denoted by

$$P(E|F)$$



A general formula for $P(E|F)$ that is valid for all events E and F. If the event F occurs, then in order for E to occur it is necessary that the actual occurrence be a point in both E and F; that is, it must be in EF. Now, since we know that F has occurred, it follows that F becomes our new (reduced) sample space and hence the probability that the event EF occurs will equal the probability of EF relative to the probability of F. That is,

$$P(E|F) = \frac{P(EF)}{P(F)}$$

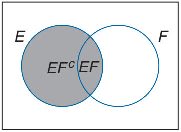
Bayes Formula

Let E and F be events. We may express E as

$$E = EF \cup EF^c$$

As EF and EF^c are clearly mutually exclusive, we have by Axiom 3 that

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)[1 - P(F)] \end{aligned}$$



Equation states that the probability of the event E is a weighted average of the conditional probability of E given that F has occurred and the conditional probability of E given that F has not occurred: Each conditional probability is given as much weight as the event it is conditioned on has of occurring.