MAT2377

Catalin Rada

Version 2009/05/05

http://www.mathstat.uottawa.ca/~crada292/

Sample spaces and events

We will deal with **random experiments** (e.g. measurements of speed/weight, number and duration of phone calls).

For any "experiment" define the **sample space** as the set of all possible outcomes. This is often denoted by the symbol $\mathcal{S}$.

A sample space can be discrete if it consists of a finite or countable infiniteset of outcomes, or continuous if it contains an interval of real numbers (e.g. $[3, \infty)$, $(1, 1.45)$).

An **event** is a collection of outcomes from the sample space $\mathcal{S}$. Events will be denoted by A, B, E_1, E_2 etc.

Examples:

- Toss a fair coin. The (discrete) sample space is $\mathcal{S} = \{\text{Head, Tail}\}$.

- Roll a die: The (discrete) s.s. is $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$ Various events:
  - Roll an even number: represent this as $\{2, 4, 6\}$.
  - Roll a prime number: $\{2, 3, 5\}$.

- Suppose we measure the weight (in grams) of a chemical sample. The (continuous) sample space can be represented by $(0, \infty)$, the pos. half line. Events:
  - sample is less than 1.5 grams: $(0, 1.5)$;
  - sample exceeds 5 grams: $(5, \infty)$;

Describe the sample space for the following experiments:

i) Each of 4 transmitted bits is classified as either in error OR not in error.

Sol: We let a be a bit in error, and we let b be a bit not in error. Then $\mathcal{S} = \{aaaa, aaab, aaba, \ldots, bbbb\}$.

ii) In the manufacturing of some recording tape, electronic testing is used to record the number of bits in error in a 351-foot reel.

Sol: $\mathcal{S} = \{0, 1, 2, 3, \ldots\}$ - the set of all nonnegative integers.

iii) A device that displays three digits is used to measure current in milliamperes.

Sol: $\{000, 001, 002, \ldots, 471, \ldots, 999\}$, the set of all vectors with 3 entries, where each entry could be one of the digits in the set: $\{0, 1, 2, \ldots, 9\}$

iv) A biologist is studying certain animals. An animal is captured and weighted. The set of all possible weights $\mathcal{S}$ is given by $\mathcal{S} = (0, \infty)$ (weight in grams). It is continuous! The biologist could also classify an animal by sex and weight. In this case the s.s. is $\mathcal{S} = \{(a, b) | a = F$ or $M, 0 < b < \infty\}$.

For any events A and B in $\mathcal{S}$:

- The Union of A and B is all outcomes from $\mathcal{S}$ in either A or B; write $A \cup B$.

- The Intersection of A and B is all outcomes in both A and B; write $A \cap B$.

- The Complement of A is all outcomes in $\mathcal{S}$ that are not in A. Write A^c OR $A^{'}$ for this.

- If A and B have no outcomes in common, they are mutually exclusive; write $A \cap B = \emptyset$ (the empty set). In particular, A and A^c are mutually exclusive.

- The symbol $\in$ means belongs to, e.g. $3.234 \in$ **R**.

- In a process, the length and width, denoted by X and Y, respectively, of each object in the process are evaluated. Suppose A is the event $47 < X < 89$, B is the event $11 < Y < 19$, C is the event that a length meets the customer request. Shade A, $B \cap A$, $B \cup A^{'}$, $B \cup A$, $C \cap B \cap A$.

Examples:

- Roll a die. Let $A = \{2, 3, 5\}$ (a prime number) and $B = \{3, 6\}$ (multiple of 3). Then $A \cup B = \{2, 3, 5, 6\}$, $A \cap B = \{3\}$ and $A^c = \{1, 4, 6\} = A'$.

- Samples of plastic are analyzed for scratch and shock resistance.

		shock resistance	
		high	low
scratch resistance	high	40	4
	low	1	5

If A is the event that a sample has high shock resistance and B is the event that a sample has high scratch resistance, then $A \cap B$ consists of 40 samples.

Counting techniques

We now consider some basic combinatorial results which make later some probabilities easier to calculate.

A **two-stage Procedure** procedure can be modeled as having k bags, with m_1 items in the first bag,. . . , m_k items in k-th bag.

The first stage consists of picking a bag, and the second stage consists of drawing an item out of that bag.

But this is equivalent to picking one of the $m_1 + m_2 + \cdots + m_k$ total number of items.

If all the bags have the same number $m_1 = \cdots = m_k = n$ then there are kn items in total, and this is the total number of ways the two-stage procedure can happen.

Examples

- How many ways can I roll a die and then draw a card from a shuffled 52-card pack?
  - There are 6 ways the first step can turn out, and for each of these (the stages are independent in fact) there are 52 ways to draw the card. Thus there are $6 \times 52 = 312$ ways this can turn out.

- How many ways can I draw out two tickets numbered 1 to 100 from a bag: first right-hand and then left hand?
  - There are 100 ways to pick the first number; for *each of these* there are 99 ways to pick the second number (*actual choices* depend on first, but *number of choices* is the same regardless). Thus 9900 ways.

Any number of stages

This leads us to general multi-stage procedures where at each stage

- the actual choices depend on the outcomes of earlier stages but
- the *number* of choices is the same regardless.

If we have a k-stage process like this, where there are n_1 possibilities for stage 1, regardless of the 1st outcome there are n_2 possibilities at stage 2, etc., up to n_k choices at stage k, then there are

$$n_1 n_2 \cdots n_k$$

total ways the process can turn out.

Ordered samples

Suppose we have items numbered $1, 2, \ldots, n$ and are we drawing an ordered sample of size r of them. Here $1, 5 \neq 5, 1$

- **with replacement**;
- **without replacement**.

How many ways can each case turn out?

Each is an r-stage procedure where *actual choices* at each stage (may) depend on earlier choices but *number of choices* does not.

With Replacement

If we replace each "number" after it is drawn, then every draw is the same and there are n ways it can turn out.

So according to our earlier result (page 9) there are

$$\underbrace{n\,n\cdots n}_{r \text{ factors}} = n^r$$

ways to choose an ordered sample of size r with replacement from $\{1, 2, \ldots, n\}$.

Without Replacement

If we **do not** replace each "number" after it is drawn, then choices for second draw depend on first draw but regardless there are $(n-1)$ choices.

Also, whatever the first two draws, there are $(n-2)$ ways to draw the third number, etc.

Thus there are

$$\underbrace{n(n-1)\cdots(n-r+1)}_{r \text{ factors}} = P_r^n \quad \text{(common calculator symbol)}$$

ways to choose an ordered sample of size $r \leq n$ without replacement from $\{1, 2, \ldots, n\}$.

Factorial notation

Writing $n! = n(n-1)(n-2)\cdots 1$, for positive integer n, we have

- when $r = n$, $P_r^n = n!$, and the ordered selection is called a **permutation**; WHY?
- when $r < n$, we can write

$$P_r^n = \frac{n(n-1)\cdots(n-r+1)}{}\frac{(n-r)\cdots 1}{(n-r)\cdots 1} = \frac{n!}{(n-r)!}\,. \qquad (1)$$

If *by convention* we take $0! = 1$ then equality (1) is also true for $r = n$.

Examples:

1. How many different ways can 6 balls be drawn *in order* without replacement from balls numbered 1 to 44?

Answer: $P_6^{44} = 44 \times 43 \times 42 \times 41 \times 40 \times 39 = 5,082,517,440.$

This is the number of ways the actual drawing of the balls could occur (e.g. watching the draw on TV).

2. Having digits $\{0, 1, \ldots, 9\}$, how many 6-digits PIN codes can you create

- if digits may be repeated: $10 \times 10 \times 10 \times 10 \times 10 \times 10 = 10^6 = 1,000,000.$

- if digits may not be repeated: $P_6^{10} = 10 \times 9 \times 8 \times 7 \times 6 \times 5.$

3. A window opener has a code determined by the down or up setting of 11 switches. If the sample space is the set of all possible codes, find how many outcomes are in the s.s.

Sol: $2 \times 2 \times \cdots \times 2 = 2^{11}$.

4. In a factory, an object is created by machining, polishing and painting. There are 3 machine devices, 4 polishing devices and 3 painting devices. How many different routings (consisting of machining, followed by polishing, and followed by painting) for an object are possible?

Sol: $3 \times 4 \times 3 = 36$

Unordered sample

Consider sampling without replacement. Now suppose that we cannot distinguish between different ordered samples that contain the same elements (e.g. looking up Lotto results in the newspaper). Denote the (as yet unknown) number of unordered samples by C_r^n.

We can deduce this by noting that the following two processes are equivalent:

- Take an ordered sample of size r.
- Take an unordered sample of size r **and then** rearrange (permute) the numbers.

There are P_r^n ways to perform the first procedure.

The second is a two-stage process. There are C_r^n ways to do the first stage. For each of these there are $r!$ ways to rearrange/permute the r numbers.

Thus

$$P_r^n = C_r^n \times r! \quad \Rightarrow \quad C_r^n = P_r^n / r! = \frac{n!}{(n-r)!r!} = \binom{n}{r}.$$

This last notation is common in many texts, and is also called a **binomial coefficient**. Read as "n-choose-r". In other words: the number of subsets of size r that can be selected from a set of n elements.

Example:

How many ways can the "Lotto draw" be reported in the newspaper (where they are always reported in increasing order)?

This is the same as the number of *unordered samples* (different reorderings of same 6 numbers are indistinguishable), and so

Answer:

$$C_6^{44} = \binom{44}{6} = \frac{44 \times 43 \times 42 \times 41 \times 40 \times 39}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{5,082,517,440}{720} = 7,059,052\,.$$

Another example: In an electronic circuit there are 12 different places (locations) that can contain chips.

a) If 5 different types of chips are to be located in the circuit, how many different circuits are possible?

b) If the 5 chips that are placed in the circuit are of the same type determine how many different circuits are possible?

Sol: a) $P_5^{12} = \frac{12!}{7!} = 95040$ possible circuits. Why? The 1st chip can be put in any of the 12 places/locations; the second chip can be put in any of the 11 left places/locations, etc.... then multiply!!!

b) Every subset of 5 places/locations chosen from our 12 different locations gives us a different circuit, so: $\binom{12}{5}$, hence $\frac{12!}{5!7!} = \frac{8\times 9\times 10\times 11\times 12}{1\times 2\times 3\times 4\times 5} = \frac{8\times 9\times 11\times 12}{3\times 4} = \frac{8\times 9\times 11}{1} = 792.$

One more example: By definition a superbyte is a sequence of 8 superbits, and every superbit is either blue or red. How many different superbytes are possible? If the first superbit of a superbyte is determined from the other 7 superbits, find how many different superbytes are possible?

Sol: $2 \times 2 \times \cdots \times 2 = 2^8$; the answer for the second question is 2^7.