EC381/MN308
Probability and Some Statistics

Yannis Paschalidis

yannisp@bu.edu, http://ionia.bu.edu/

Dept. of Manufacturing Engineering

Dept. of Electrical and Computer Engineering

Center for Information and Systems Engineering

Lecture 4 - Outline

1. Counting.
2. Reliability theory.

1.5 Principles of Counting
for problems with uniform probability spaces

In many experiments with finite sample spaces S, all outcomes in the sample space S are "equally likely"

Estimating probabilities requires knowing how many outcomes

$$P[A] = \frac{\text{outcomes in } A}{\text{total number of outcomes}}$$

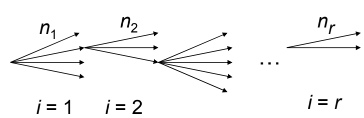
This requires counting (combinatorics)

Not central focus of course

1. Number of outcomes of sequential experiments

If subexperiment A_1 has n_1 possible outcomes and subexperiment A_2 has n_2 possible outcomes then there are $n_1 n_2$ possible outcomes when both experiments are performed

For r subexperiments with n_i outcomes for subexperiment i, if all choices can be made freely, the total number of outcomes is: $\boxed{n_1\, n_2 \ldots n_r}$



Examples

- 3 pairs of shoes, 2 pairs of pants, 4 shirts:
  Number of outfits = 3x2x4
- Number of license plates with 3 digits:
  3 stages, 10 choices per stage: $10^3 = 1000$

  2 4 3

  Number of license plates with 3 distinct digits
  3 stages: 10 choices at first stage, 9 at 2nd, 8 at 3rd
  Sampling without replacement…
- Card deck: Number of 5-card poker (7 & up) hands:
  5 stages: 8 x 7 x 6 x 5 x 4

2. Sampling

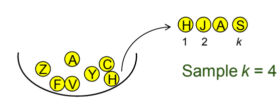
Given a population of n distinguishable elements $\{1, \ldots, n\}$, k elements are sampled.

How many distinguishable samples are possible (if we care about order and if we don't)?

The sampling may be done with or without replacement.

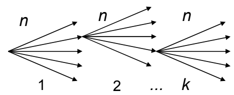
If sampling is done with replacement, elements of the sample may be repeated.

Example



Population $n = 26$

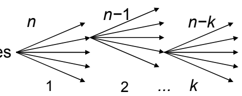
A. Sampling with replacement, order dependent



# of possibilities = n^k

B. Sampling without replacement, order dependent

$n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$ possibilities



Number of possible subsets of *k* elements out of set of *n* is $\frac{n!}{(n-k)!}$ permutations

If $k = n$: $n!$ possibilities

Number of ways of ordering a set {1, …, *n*} with *n* elements is *n*! permutations

C. Sampling without replacement, order independent

Number of ways of picking *k* elements out of a set of *n* elements:

=Binomial coefficients $C_k^n \equiv \binom{n}{k} = \frac{n!}{k!(n-k)!}$

Number of possible subsets of k elements out of set of n is C_k^n combinations

D. Sampling with replacement, order independent

Also binomial coefficient, but harder formula…

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

Binomial Theorem $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Helps in computing many identities $x = 1, y = 1 \Rightarrow 2^n = \sum_{k=0}^{n} \binom{n}{k}$

E. Partitioning

Consider an n-element set and nonnegative integers $n_0, \ldots, n_{m-1}$ such that $n_0+\cdots+n_{m-1}=n$. Partition the set into m disjoint subsets as follows:

n_0 *elements*	n_1 *elements*	…	n_{m-1} *elements*

We can do that in

$$\binom{n}{n_0, n_1, \ldots, n_{m-1}} = \frac{n!}{n_0! n_1! \cdots n_{m-1}!}$$ ways. (Multinomial Coefficient)

Examples

- Number of hexadecimal (base-16) numbers with 4 characters: 16^4

  Sample (2)(12)(7)(15)

- Number of 8-bit binary symbols with 3 ones: $\binom{8}{3}$

  Sample 01000011

- Birthday Paradox: Consider a group of r people. The probability they have distinct birthdays is:

$$\frac{\frac{365!}{(365-r)!}}{365^r} = \frac{365(365-1)\cdots(365-r+1)}{365^r}$$
$$= \left(1-\frac{1}{365}\right)\cdots\left(1-\frac{r-1}{365}\right)$$
$$= \begin{cases} 0.883 & r = 10 \\ 0.4927 & r = 23 \\ 0.0059 & r = 60 \end{cases}$$

3. Independent Trials

- Many compound experiments are sequences of independent subexperiments
  - Repeated coin tosses
  - Multiple bits through communication channel,
  - Multiple rolls of 6 sided die
- Objective: compute probabilities of events
  - Total number of heads = *k*
  - Total number of bit errors = *k*
  - Sum of numbers = *k*

A. Subexperiments with binary outcomes

Subexperiment:
- Outcome 1: Success; probability *p*
- Outcome 2: Failure; probability 1 − *p*

Experiment: *n* subexperiments
- Outcome: a sequence of length n of failures and successes SSFS...F
- Event *E*: *k* failures in *n* attempts

Probability of *E* = *k* failures in *n* attempts
- Probability of one outcome with k failures: $(1-p)^k p^{n-k}$
- Since subexperiment events are independent &
- number of outcomes with *k* failures = combinations of *k* out of *n*,

$P[E] = P[k$ failures in n attempts$] = \binom{n}{k}(1-p)^k p^{n-k}$

Example 1

Communication channel has single bit error rate of p, independent across each bit

Redundant coding of 8 bits into 9 bits allows detection of single bit errors in 8 bit words

- Parity check: makes sure that checksum of all bits is even

What is probability that a signal with errors is not detected as an error?

- What is the probability that the number of errors in 9 bits is even?
- P[2 errors] + P[4 errors] + P[6 errors] + P[8 errors]

$$P[k \text{ errors}] = \binom{8}{k} (p)^k (1-p)^{8-k}$$

Example 2 Quiz 1.9

100 bit packets, single bit prob of error 0.01, can correct 3 bit errors

$$P[k \text{ errors}] = \binom{100}{k} (0.01)^k (0.99)^{100-k}$$

P[uncorrected error] =

1 – P[0 errors] – P[1 error] – P[2 errors] – P[3 errors]

B. Subexperiments with N-ary outcomes

Each subexperiment has N-ary outcomes with probabilities p_k, $k = 1, \cdots, N$. There are M subexperiments (independent trials).

Probability that one gets $n_1, \cdots, n_N$ outcomes of each type?

- Outcomes : length M sequences
- Probability of each sequence with $n_1, \cdots, n_N$ outcomes of each type:

$$p_1^{n_1} p_2^{n_2} \dots p_N^{n_N}$$

- Number of sequences with $n_1, n_2, \ldots, n_N$ outcomes of each type:

$$\frac{M!}{n_1! n_2! \cdots n_N!}$$

Example $M = 8$

$N = 6$

$p_k = 1/6$

$$P[(n_1, \ldots, n_N) \text{ outcomes}] = \frac{M!}{n_1! n_2! \cdots n_N!} p_1^{n_1} \cdots p_N^{n_N}$$

Example

Calls arriving into a bank's call center are either concerning Checking accounts (C), Savings accounts (S) or Investment accounts (I).

P[C]= 0.4, P[S]=0.4, P[I]=0.2.

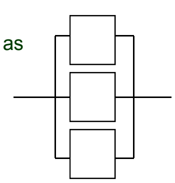
$S_{c,s,i}$ = {Observe c type C, s type S, and i type I calls out of 1000}

$$P[S_{c,s,i}] = \binom{1000}{c, s, i} (0.4)^c (0.4)^s (0.2)^i$$

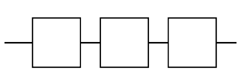
Application: Reliability Theory for Interconnected Systems

Assume independent failures of individual subcomponents

Parallel connection: System works as long as one of subsystems works (redundancy)



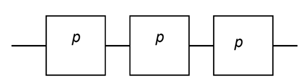
Series connection: System works as long as all subsystems work



Probability of System Failure

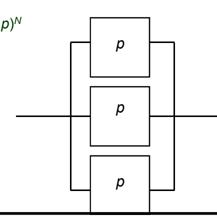
For N subsystems in series, each of them has probability p of working correctly

- E = event system works; E_i = event subsystem works
- $P[E] = P[E_1]\, P[E_2] \ldots P[E_N] = p^N$



For N components in parallel:

- $P[E^c] = P[E_1^c]\, P[E_2^c] \ldots P[E_N^c] = (1-p)^N$
- $P[E] = 1 - P[E^c] = 1 - (1-p)^N$



Example 1.44

$p = 0.9$. E = system works; E_k = subsystem k works

$E = (E_1 \cap E_2) \cup (E_3 \cap E_4)$; $P[E_1 \cap E_2] = p^2 = P[E_3 \cap E_4]$

Interpret as parallel connection of 2 subsystems, each with success probability $p^2 \rightarrow P[E] = 1 - (1 - p^2)^2$