Counting 2

Ruben Zamar
Department of Statistics
UBC

January 8, 2019

BASIC PRINCIPLE OF COUNTING

- Experiment has k steps

  Step 1 has n_1 possible outcomes
  Step 2 has n_2 possible outcomes
  \- -
  \- -
  Step k has n_k possible outcomes

- Number of possible outcomes for the experiment

$$n_1 \times n_2 \times \cdots \times n_k$$

PERMUTATION

- $g : \{1, 2, ..., n\} \to \{1, 2, ..., n\}$ one-to-one and onto

  Example: let $n = 4$

$$\begin{array}{ccc} 1 & \to & 3 \\ 2 & \to & 4 \\ 3 & \to & 1 \\ 4 & \to & 2 \end{array}$$

- How many permutations can be defined?

$$n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 = n!$$

Combinations

- Split a set of n objects into 2 subsets of sizes m and $n-m$, respectively.

Ex: $\{1,2,3,4,5\}$ is split into two subsets $\{1,3,5\}$ and $\{2,4\}$
In this case $n=5$ and $m=3$

How many splits (combinations) can be formed?

1,2,3	4,5
1,2,4	3,5
1,2,5	3,4
1,3,4	2,5
1,3,5	2,4

1,4,5	2,3
2,3,4	1,5
2,3,5	1,4
2,4,5	1,3
3,4,5	1,2

In this case there are **10 possible splits**.

PERMUTATIONS GENERATE SPLITS

Each permutation produces an split: **the first** m and **the last** $m - n$

$$\overbrace{i_1, i_2, i_3}^{m=3}, \quad \overbrace{i_4, i_5}^{n-m=2}$$

Example:

$$\overbrace{1, 5, 4}, \quad \overbrace{2, 3}$$

But there is duplication, for example: $\overbrace{5, 4, 1}, \quad \overbrace{3, 2}$

ACCOUNTING FOR DUPLICATIONS

Divide by the number of permutations in each split:

$$\underbrace{{}_nC_m}_{\text{"}n\text{ choose }m\text{"}} = \left(\begin{array}{c} n \\ \\ m \end{array}\right) = \frac{\overbrace{n!}^{\#\text{ permutations}}}{\underbrace{m!}_{\#\text{ permutations}}\ \underbrace{(n-m)!}_{\#\text{ permutations}}}$$

For example

$$_5C_3 = \left(\begin{array}{c} 5 \\ \\ 3 \end{array}\right) = \frac{5!}{3!2!} = 10$$

WHAT IF THE ORDER IS IMPORTANT?

If the order is important (in the first split) we must multiply back by $m!$

$$ {}_nP_m \quad = \quad m! \begin{pmatrix} n \\ \\ m \end{pmatrix} = \frac{n!}{(n-m)!} $$

EXAMPLE: LOTTERY 6/49 (continued)

- Matching exactly x numbers and missing the other $6 - x$ numbers
- **Mind Experiment:** Consider a box with 50 balls numbered 0 to 49.

  Six of these balls are labeled "W" (your six chosen numbers)

  The remaining 44 are labeled "L" (the non chosen numbers)
- Formula:

$$p(x) = \frac{\binom{6}{x}\binom{44}{6-x}}{\binom{50}{6}}$$

RESULTS

x	$p(x)$
0	0.44423
1	0.41005
2	0.12814
3	0.01667
4	0.00089
5	0.00002
6	0.00000

In fact, $p(6) = 6.292989\text{e-}08$

APPLICATION TO CRYPTOGRAPHY

Let

$$\mathcal{A} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$$

be a given alphabet and let W be text of length N over $\mathcal{A}$.

- Assume that each N-character text, W, is equally likely.
- α_i is a *blank* if α_i doesn't appear in W.

1. Calculate the probability that the text W contains at least one blank.
2. Calculate the probability [denoted $P(N, n, 0)$] that the text W contains no blanks.
3. Calculate the probability [denoted $P(N, n, b)$] that the text W contains b blanks, $b = 1, ..., n - 1$.

APPLICATION TO CRYPTOGRAPHY

Let

$$A_i = \{\alpha_i \ \text{ is a blank}\}$$

$$p_1 = P(A_i) = \frac{(n-1)^N}{n^N} = \left(1 - \frac{1}{n}\right)^N$$

$$p_2 = P(A_{i_1} \cap A_{i_2}) = \frac{(n-2)^N}{n^N} = \left(1 - \frac{2}{n}\right)^N, \ \text{for } i_1 < i_2$$

and in general

$$p_k = P(A_{i_1} \cap \cdots \cap A_{i_k}) = \left(1 - \frac{k}{n}\right)^N, \ \text{for } i_1 < \cdots < i_k$$

AT LEAST ONE BLANK

$$
\begin{aligned}
P(\text{At least one blank}) &= P(A_1 \cup \cdots \cup A_n) \\
&= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} p_k \\
&= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^N
\end{aligned}
$$

ZERO BLANKS

$$
\begin{aligned}
P\left(\text{Zero blanks}\right) &= P\left(N, n, 0\right) = 1 - P\left(A_1 \cup \cdots \cup A_n\right) \\
&= 1 - \sum_{k=1}^{n} \left(-1\right)^{k-1} \binom{n}{k} \left(1 - \frac{k}{n}\right)^N \\
&= 1 + \sum_{k=1}^{n} \left(-1\right)^{k} \binom{n}{k} \left(1 - \frac{k}{n}\right)^N \\
&= \sum_{k=0}^{n} \left(-1\right)^{k} \binom{n}{k} \left(1 - \frac{k}{n}\right)^N
\end{aligned}
$$

TEXT HAS b BLANKS

It can be shown that, for all $b = 0, 1, ..., n-1$

$$\begin{aligned} P(N, n, b) &= \left(\begin{array}{c} n \\ \\ b \end{array}\right) \left(1 - \frac{b}{n}\right)^N P(N, n-b, 0) \\ &= \left(\begin{array}{c} n \\ \\ b \end{array}\right) \left(1 - \frac{b}{n}\right)^N \sum_{k=0}^{n-b} (-1)^k \left(\begin{array}{c} n-b \\ \\ k \end{array}\right) \left(1 - \frac{k}{n-b}\right)^N \end{aligned}$$

NUMERICAL CALCULATIONS

Suppose (for simplicity) that the Alphabet $\mathcal{A}$ has $n = 30$ symbols and the text W has length $N = 90$. In this case,

b	$P(N, n, b)$
0	0.20645
1	0.36527
2	0.27489
3	0.11637
4	0.03089
5	0.00543
6	0.00065
7	0.00005
8	0.00000

N = Text length = 90
n = Alphabet size = 30