Statistics for Communications and Security

Cryptography

School of Engineering and Technology
CQUniversity Australia

Prepared by Steven Gordon on 19 Feb 2020,
statistics.tex, r1791

Contents

Properties of Exponentials and Logarithms

$$n^x \times n^y = n^{x+y}$$

$$\frac{n^x}{n^y} = n^{x-y}$$

$$\log_n (x \times y) = \log_n(x) + \log_n(y)$$

$$\log_n \left(\frac{x}{y}\right) = \log_n(x) - \log_n(y)$$

Properties of Exponentials (example)

Properties can be applied to simplify calculations:

$$\begin{aligned} 2^{12} &= 2^{2+10} \\ &= 2^2 \times 2^{10} \\ &= 4 \times 1024 \\ &= 4096 \end{aligned}$$

With this property of exponentials, if you can remember the values of 2^1 to 2^{10} then you can approximate most values of 2^b that you come across in communications and security. Table 5 gives the exact or approximate decimal value for b-bit numbers.

Useful Exact and Approximate Values in Binary

Exponent, b	2^b	
(bits)	Exact Value	Approx. Value
0	1	-
1	2	-
2	4	-
3	8	-
4	16	-
5	32	-
6	64	-
7	128	-
8	256	-
9	512	-
10	1,024	$1{,}000 = 10^3$
11	-	2,000
12	-	4,000
13	-	8,000
14	-	16,000
	...	
19	-	512,000
20	-	$1{,}000{,}000 = 10^6$
21	-	2×10^6
22	-	4×10^6
23	-	8×10^6
	...	
29	-	512×10^6
30	-	10^9
31	-	2×10^9
32	-	4×10^9
33	-	8×10^9
	...	
39	-	512×10^9
40	-	10^{12}
50	-	10^{15}
60	-	10^{18}
70	-	10^{21}
$x \times 10$	-	10^{3x}

Properties of Exponentials with Binary Values (example)

Properties and approximations can be used to perform large calculations:

$$\begin{aligned}\frac{2^{128}}{2^{100}} &= 2^{128-100} \\ &= 2^{28} \\ &= 2^{8} \times 2^{20} \\ &\approx 256 \times 10^{6} \\ &\approx 10^{8}\end{aligned}$$

Properties of Logarithms (example)

The number of bits needed to represent a decimal number can be found using logarithms:

$$\begin{aligned} \log_2(20,000) &= \log_2(20 \times 10^3) \\ &= \log_2(20) + \log_2(10^3) \\ &\approx 4 + 10 \\ &\approx 14 \end{aligned}$$

Contents

Number of Binary Values (definition)

Given an n-bit number, there are 2^n possible values.

Number of Sequence Numbers (example)

Consider a sliding-window flow control protocol that uses an 16-bit sequence number. There are $2^{16} = 65,536$ possible values of the sequence number, ranging from 0 to 65,535 (after which it wraps back to 0).

Number of IP Addresses (example)

An IP address is a 32-bit value. There are 2^{32} or approximately 4×10^9 possible IP addresses.

Number of Keys (example)

If choosing a 128-bit encryption key randomly, then there are 2^{128} possible values of the key.

Fixed Length Sequences (definition)

Given a set of n items, there are n^k possible k-item sequences, assuming repetition is allowed.

Sequences of PINs (example)

A user chooses a 4-digit PIN for a bank card. As there are 10 possible digits, there are 10^4 possible PINs to choose from.

Sequences of Keyboard Characters (example)

A standard keyboard includes 94 printable characters (a–z, A–Z, 0–9, and 32 punctuation characters). If a user must select a password of length 8, then there are 94^8 possible passwords that can be selected.

Pigeonhole Principle (definition)

If n objects are distributed over m places, and if $n > m$, then some places receive at least two objects.

Pigeonhole Principle on Balls (example)

There are 20 balls to be placed in 5 boxes. At least one box will have at least two balls. If the balls are distributed in a uniform random manner among the boxes, then on average there will be 4 balls in each box.

Pigeonhole Principle on Hash Functions (example)

A hash function takes a 100-bit input value and produces a 64-bit hash value. There are 2^{100} possible inputs distributed to 2^{64} possible hash values. Therefore at least some input values will map to the same hash value, that is, a collision occurs. If the hash function distributes the input values in a uniform random manner, then on average, there will be $\frac{2^{100}}{2^{64}} \approx 6.4 \times 10^{10}$ different input values mapping to the same hash value.

Contents

Factorial (definition)

There are $n!$ different ways of arranging n distinct objects into a sequence.

Factorial and Balls (example)

Consider four coloured balls: Red, Green, Blue and Yellow. There are $4! = 24$ arrangements (or permutations) of those balls:

RGBY, RGYB, RBGY, RBYG, RYGB, RYBG,
GRBY, GRYB, GBRY, GBYR, GYRB, GYBR,
BRGY, BRYG, BGRY, BGYR, BYRG, BYGR,
YRGB, YRBG, YGRB, YGBR, YBRG, YBGR

Factorial and English Letters (example)

The English alphabetic has 26 letters, a–z. There are $26! \approx 4 \times 10^{26}$ ways to arrange those 26 letters.

Factorial and Plaintext Messages (example)

An encryption algorithm takes a 64-bit plaintext message and a key as input and then maps that to a 64-bit ciphertext message as output. There are $2^{64} \approx 1.6 \times 10^{19}$ possible input plaintext messages. There are $2^{64}! \approx 10^{10^{88}}$ different reversible mappings from plaintext to ciphertext, i.e. $2^{64}!$ possible keys.

Combinations (definition)

The number of combinations of items when selecting k at a time from a set of n items, assuming repetition is not allowed and order doesn't matter, is:

$$\frac{n!}{k!\,(n-k)!}$$

Number of Pairs (definition)

The number of pairs of items in a set of n items, assuming repetition is not allowed and order doesn't matter, is:

$$\frac{n(n-1)}{2}$$

Pairs of Coloured Balls (example)

There are four coloured balls: Red, Green, Blue and Yellow. The number of different coloured pairs of balls is $4 \times 3/2 = 6$. They are: `RG`, `RB`, `RY`, `GB`, `GY`, `BY`. Repetitions are not allowed (as they won't produce different coloured pairs), meaning `RR` is not a valid pair. Ordering doesn't matter, meaning `RG` is the same as `GR`.

Pairs of Network Devices (example)

A computer network has 10 devices. The number of links needed to create a full-mesh topology is $10 \times 9/2 = 45$.

Pairs of Key Sharers (example)

There are 50 users in a system, and each user shares a single secret key with every other user. The number of keys in the system is $50 \times 49/2 = 1,225$.

Contents

Probability of Selecting a Value (definition)

Probability of randomly selecting a specific value from a set of n values is $1/n$.

Probability of Selecting Coloured Ball (example)

There are five coloured balls in a box: red, green, blue, yellow and black. The probability of selecting the yellow ball is $1/5$.

Probability of Selecting Backoff Value (example)

IEEE 802.11 (WiFi) involves a station selecting a random backoff from 0 to 15. The probability of selecting 5 is 1/16.

Total Expectation (definition)

For a set of n events which are mutually exclusive and exhaustive, where for event i the expected value is E_i given probability P_i, then the total expected value is:

$$E = \sum_{i=1}^{n} E_i P_i$$

Total Expectation of Packet Delay (example)

Average packet delay for packets in a network is 100 ms along path 1 and 150 ms along path 2. Packets take path 1 30% of the time, and take path 2 70% of the time. The average packet delay across both paths is:
$100 \times 0.3 + 150 \times 0.7 = 135$ ms.

Total Expectation of Password Length (example)

In a network with 1,000 users, 150 users choose a 6-character password, 500 users choose a 7-character password, 250 users choose 9-character password and 100 users choose a 10-character password. The average password length is 7.65 characters.

Number of Attempts (definition)

If randomly selecting values from a set of n values, then the number of attempts needed to select a particular value is:

best case: 1

worst case: n

average case: $n/2$

Number of Attempts in Choosing Number (example)

One person has chosen a random number between 1 and 10. Another person attempts to guess the random number. The best case is that they guess the chosen number on the first attempt. The worst case is that they try all other numbers before finally getting the correct number, that is 10 attempts. If the process is repeated 1000 times (that is, one person chooses a random number, the other guesses, then the person chooses another random number, and the other guesses again, and so on), then on average 10% of time it will take 1 attempt (best case), 10% of the time it will take 2 attempts, 10% of the time it will take 3 attempts, ..., and 10% of the time it will take 10 attempts (worst case). The average number of attempts is therefore 5.

Number of Attempts in Choosing Key (example)

A user has chosen a random 128-bit encryption key. There are 2^{128} possible keys. It takes an attacker on average $2^{128}/2 = 2^{127}$ attempts to find the key. If instead a 129-bit encryption key was used, then the attacker would take on average $2^{129}/2 = 2^{128}$ attempts. (Increasing the key length by 1 bit doubles the number of attempts required by the attacker to guess the key).

Contents

Birthday Paradox (definition)

Given n random numbers selected from the range 1 to d, the probability that at least two numbers are the same is:

$$p(n; d) \approx 1 - \left(\frac{d-1}{d}\right)^{n(n-1)/2}$$

Two People Have Same Birthday (example)

Given a group of 10 people, the probability of at least two people have the same birth date (not year) is:

$$p(10; 365) \approx 1 - \left(\frac{364}{365}\right)^{10(9)/2} = 11.6\%$$

Defintion 40 can be re-arranged to find the number of values needed to obtain a specified probability that at least two numbers are the same:

$$n(p; d) \approx \sqrt{2d \ln\left(\frac{1}{1-p}\right)}$$

Group Size for Birthday Matching (example)

How many people in a group are needed such that the probability of at least two of them having the same birth date is 50%?

$$n(0.5; 365) \approx \sqrt{2 \times 365 \times \ln\left(\frac{1}{1-0.5}\right)} = 22.49$$

So 23 people in a group means there is 50% chance that at least two have the same birth date.

Group Size for Hash Collision (example)

Given a hash function that outputs a 64-bit hash value, how many attempts are need to give a 50% chance of a collision?

$$\begin{aligned} n(0.5; 2^{64}) &\approx \sqrt{2 \times 2^{64} \times \ln\left(\frac{1}{1-0.5}\right)} \\ &\approx \sqrt{2^{64}} \\ &= 2^{32} \end{aligned}$$

Following Example 43, the number of attempts to produce a collision when using an n-bit hash function is approximately $2^{n/2}$.