Stats for Engineers: Lecture 3

Conditional probability

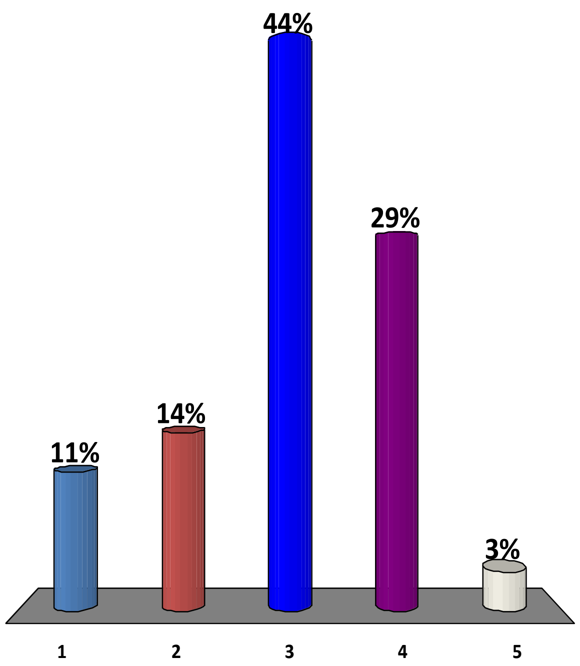
Suppose there are three cards:

A *red card* that is red on both sides,
A *white card* that is white on both sides, and
A *mixed card* that is red on one side and white on the other.

All the cards are placed into a hat and one is pulled at random and placed on a table.

If the side facing up is red, what is the probability that the other side is also red?

1. 1/6
2. 1/3
3. 1/2
4. ✓ 2/3
5. 5/6



Conditional probability

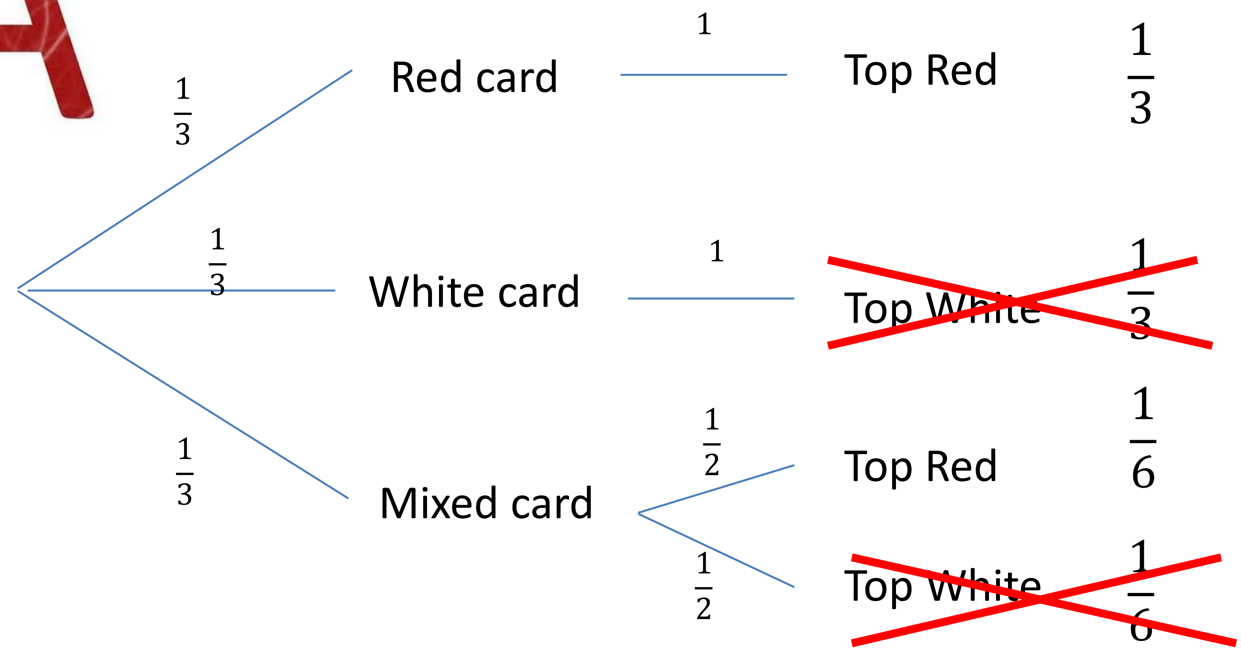
Suppose there are three cards:

A *red card* that is red on both sides,
A *white card* that is white on both sides, and
A *mixed card* that is red on one side and white on the other.

All the cards are placed into a hat and one is pulled at random and placed on a table.

If the side facing up is red, what is the probability that the other side is also red?

Probability tree



Let R=red card, TR = top red.

$$P(R|TR) = \frac{P(R \cap TR)}{P(TR)}$$

$$= \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}}$$

$$= \frac{2}{3}$$

Conditional probability

Suppose there are three cards:

A *red card* that is red on both sides,
A *white card* that is white on both sides, and
A *mixed card* that is red on one side and white on the other.

All the cards are placed into a hat and one is pulled at random and placed on a table.

If the side facing up is red, what is the probability that the other side is also red?

Let R=red card, W = white card, M = mixed card. Let TR = top is a red face.

For a random draw P(R)=P(W)=P(M)=1/3.

Total probability rule: $$P(TR) = P(TR|R)P(R) + P(TR|M)P(M)$$

$$= 1 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} \quad = \frac{1}{2}$$

The probability we want is P(R|TR) since having the red card is the only way for the other side also to be red.

This is

$$P(R|TR) = \frac{P(TR|R)P(R)}{P(TR)} \quad = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} \quad = \frac{2}{3}$$

Intuition: 2/3 of the three red faces are on the red card.

Summary From Last Time

Bayes' Theorem $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ e.g. from $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Total Probability Rule: $P(B) = \sum_k P(B|A_k)P(A_k)$

Permutations - ways of ordering k items: k!

Ways of choosing *k* things from *n, irrespective of ordering:*

$$C_k^n = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Random Variables: *Discreet* and *Continuous*

Mean $\mu = E\big(f(X)\big) \equiv \langle f(X)\rangle = \sum_k f(k)P(X=k)$

Means add: $\langle aX + bY\rangle = \langle aX\rangle + \langle bY\rangle = a\langle X\rangle + b\langle Y\rangle = a\mu_X + b\mu_Y$

Mean of a product of independent random variables

If X and Y are *independent* random variables, then $P(X \cap Y) = P(X)P(Y)$

$$\langle XY \rangle = \sum_x \sum_y P(x \cap y)xy = \sum_x \sum_y P(x)P(y)xy$$

$$= \sum_x P(x)x \sum_y P(y)y$$

$$= \langle X \rangle \langle Y \rangle = \mu_X \mu_Y$$

Note: in general this is not true if the variables are not independent

Example: If I throw two dice, what is the mean value of the product of the throws?

The mean of one throw is $\mu = \sum_{k=1}^{6} kP(X = k)$

$$= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$$

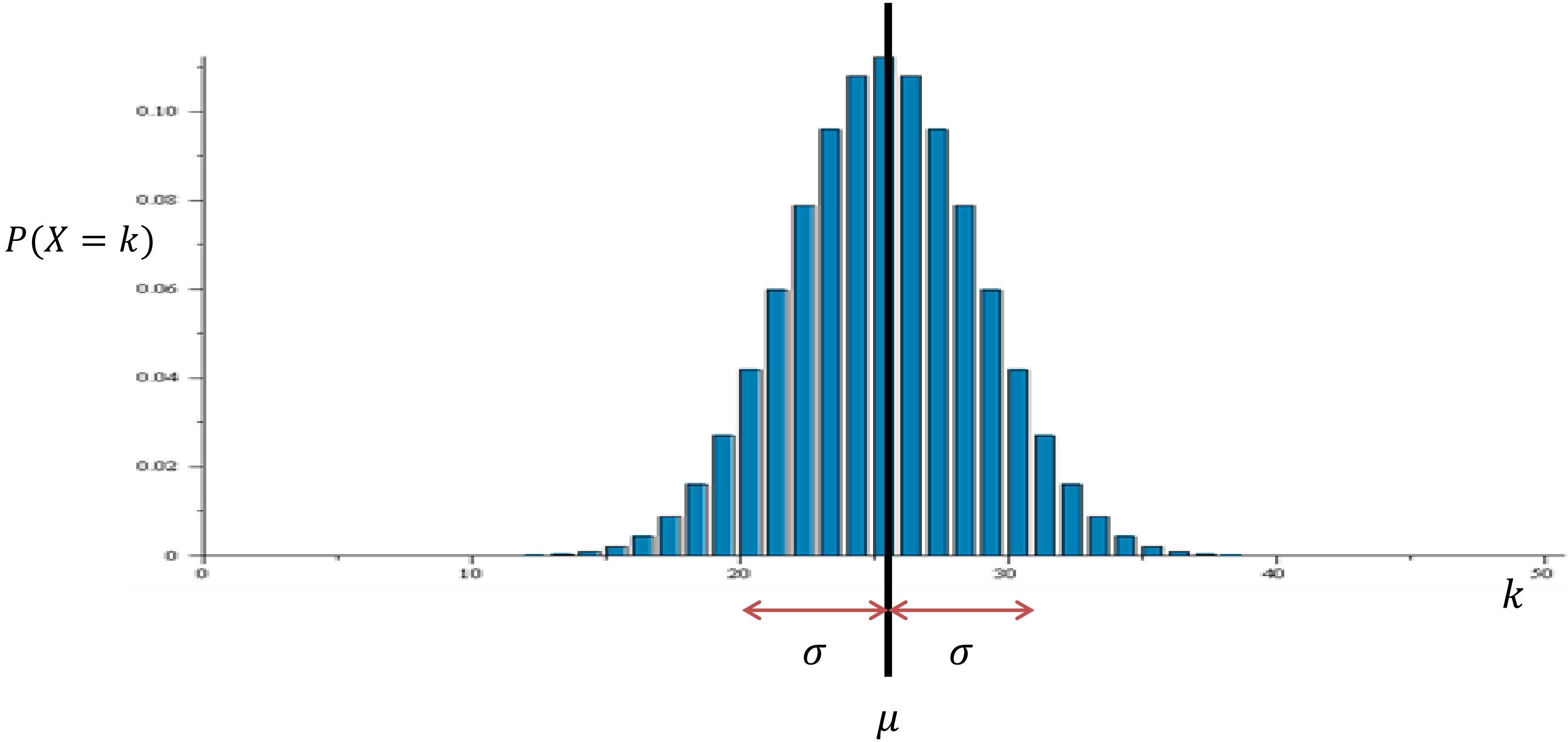
$$= (1 + 2 + 3 + 4 + 5 + 6) \times \frac{1}{6} = \frac{21}{6} = 3.5$$

Two throws are independent, so $\langle X_1 X_2 \rangle = \mu_{X_1} \mu_{X_2} = 3.5^2 = 12.25$

Variance and standard deviation of a distribution

For a random variable *X* taking values 0, 1, 2 the mean μ is a measure of the average value of a distribution, $\mu = \langle X \rangle$.

The standard deviation, σ , is a measure of how spread out the distribution is



Definition of the variance (=σ^2)

$$\sigma^2 \equiv \mathrm{var}(X) = \langle (X-\mu)^2 \rangle = \sum_k (k-\mu)^2 P(X=k)$$

Note that

$$\mu\langle X\rangle = \mu^2$$

$$\langle (X-\mu)^2 \rangle = \langle X^2 - 2\,X\,\mu + \mu^2 \rangle \quad = \langle X^2 \rangle - 2\,\langle X\mu \rangle + \mu^2$$

$$= \langle X^2 \rangle - 2\mu^2 + \mu^2$$

$$= \langle X^2 \rangle - \mu^2$$

So the variance can also be written

$$\sigma^2 = \mathrm{var}(X) = \langle X^2 \rangle - \mu^2 = \sum_k k^2 P(X=k) - \mu^2$$

This equivalent form is often easier to evaluate in practice, though can be less numerically stable (e.g. when subtracting two large numbers).

Example: what is the mean and standard deviation of the result of a dice throw?

Answer: Let X be the random variable that is the number on the dice

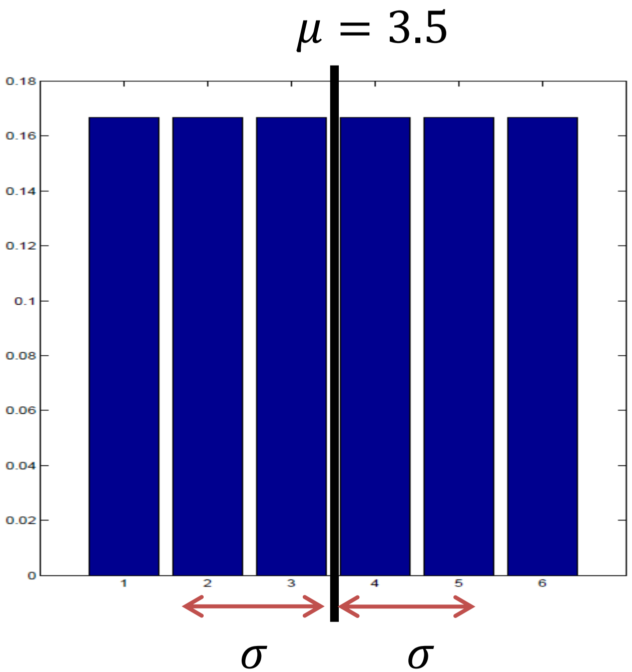
The mean is $\mu = 3.5$ as shown previously.

The variance is $\sigma^2 = \sum_{k=1}^{6} k^2 P(X = k) - \mu^2$

$$= (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) \times \frac{1}{6} - 3.5^2$$

$$= \frac{91}{6} - 3.5^2 \approx 2.917$$

Hence the standard deviation is $\sigma = \sqrt{2.917} \approx 1.71$



Sums of variances

For two *independent* (or just uncorrelated) random variables X and Y the variance of X+Y is given by the sums of the separate variances.

Why? If X has $\langle X\rangle = \mu_X$, and Y has $\langle Y\rangle = \mu_Y$, then

$$\langle X+Y\rangle = \langle X\rangle + \langle Y\rangle = \mu_X + \mu_Y.$$

Hence since var $(Z) = \langle (Z-\mu_Z)^2\rangle$, if $Z = X+Y$ then

$$\text{var}(X+Y) = \langle (X+Y-\mu_X-\mu_Y)^2\rangle = \langle [(X-\mu_X)+(Y-\mu_Y)]^2\rangle$$

$$= \langle (X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)\big(Y-\mu_y\big)\rangle$$

$$= \langle (X-\mu_X)^2\rangle + \langle (Y-\mu_Y)^2\rangle + 2\langle (X-\mu_X)\big(Y-\mu_y\big)\rangle$$

If X and Y are independent (or just uncorrelated) then

$$\langle (X-\mu_X)(Y-\mu_Y)\rangle = \langle (X-\mu_X)\rangle\langle (Y-\mu_Y)\rangle = (\mu_X-\mu_X)(\mu_Y-\mu_Y) = 0$$

Hence

$$\text{var}(X+Y) = \langle (X-\mu_X)^2\rangle + \langle (Y-\mu_Y)^2\rangle$$

$= \text{var}(X) + \text{var}(Y)$ ["Variances add"]

In general, for both discrete and continuous independent (or uncorrelated) random variables

$$\mathrm{var}(X+Y+Z+\cdots) = \mathrm{var}(X)+\mathrm{var}(Y)+\mathrm{var}(Z)+\cdots$$

Example:

The mean weight of people in England is μ=72.4kg, with standard deviation σ =15kg.

What is the mean and standard deviation of the weight of the passengers on a plane carrying 200 people?

Answer:

In reality be careful - assumption of independence unlikely to be accurate

The total weight $M = \sum_{i=1}^{200} m_i$

Since means add $\mu_M = \sum_{i=1}^{200} \langle m_i \rangle = 200 \times 72.4\mathrm{Kg} = 14480\mathrm{Kg}$

Assuming weights independent, variances also add, with $\sigma^2 = 15^2\mathrm{Kg}^2 = 225\ \mathrm{Kg}^2$

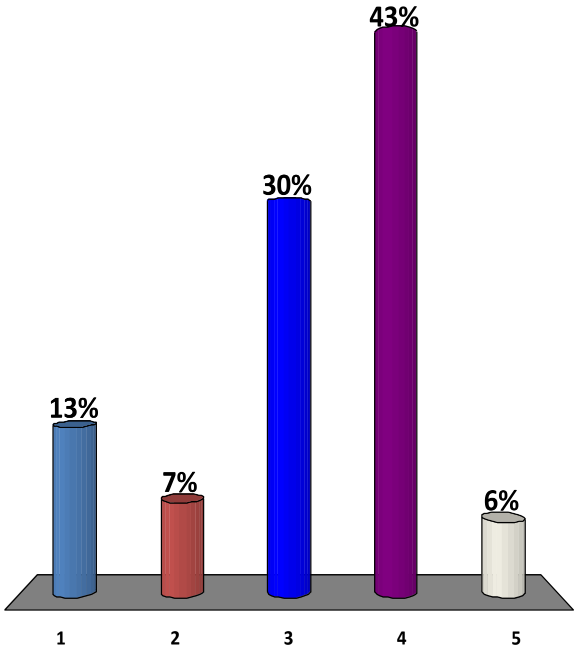
$$\sigma_M^2 = \sum_{i=1}^{200} 225\mathrm{Kg}^2 = 200 \times 225\ \mathrm{Kg}^2 = 45000\mathrm{Kg}^2 \quad \Rightarrow \quad \sigma = \sqrt{45000\mathrm{Kg}^2} \approx 212\ \mathrm{Kg}$$

Error bars

A bridge uses 100 concrete slabs, each weighing (10 ± 0.1) tonnes
[i.e. the standard deviation of each is 0.1 tonnes]

What is the total weight in tonnes of the concrete slabs?

1. 1000 ± 0.01
2. 1000 ± 0.1
3. 1000 ± 1 ✓
4. 1000 ± 10
5. 1000 ± 100



Error bars

A bridge uses 100 concrete slabs, each weighing (10 ± 0.1) tonnes [i.e. the standard deviation of each is 0.1 tonnes]

What is the total weight in tonnes of the concrete slabs?

Means add, so $\mu_{tot} = 100 \times 10 = 1000\ tonnes$

Variances add, with $\sigma^2 = 0.1^2$, so $\sigma_{tot}^2 = 100 \times 0.1^2 = 1$

Hence $M_{tot} = \left(1000 \pm \sqrt{1}\right) tonnes = (1000 \pm 1) tonnes$

Note: Error grows with the *square root* of the number: $\propto \sqrt{N}$

But the mean of the total is $\propto N$

$\Rightarrow$ *fractional* error *decreases* $\propto 1/\sqrt{N}$

Discrete Random Variables

Reminder:

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

Binomial distribution

A process with two possible outcomes, "success" and "failure" (or yes/no, etc.) is called a *Bernoulli trial.*

e.g.	coin tossing:	Heads or Tails
	quality control:	Satisfactory or Unsatisfactory
	Polling:	Agree or disagree

An experiment consists of *n* independent Bernoulli trials and *p* = probability of success for each trial. Let *X* = total number of successes in the *n* trials.

Then $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, 2, \ldots, n$.

This is called the Binomial distribution with parameters *n* and *p*, or B(*n*, *p*) for short.

X ~ B(*n*, *p*) stands for "*X* has the Binomial distribution with parameters *n* and *p*."

Situations where a Binomial might occur

1) Quality control: select n items at random; X = number found to be satisfactory.

2) Survey of n people about products A and B; X = number preferring A.

3) Telecommunications: n messages; X = number with an invalid address.

4) Number of items with some property above a threshold; e.g. X = number with height > A

Justification

"*X* = *k*" means *k* successes (each with probability *p*) and *n*-*k* failures (each with probability 1-*p*).

Suppose for the moment all the successes come first. Assuming independence

$$\text{probability} = \underbrace{p \times p \times p \ldots \times p}_{k \text{ successes: } p^k} \times \underbrace{(1-p) \times (1-p) \times \cdots \times (1-p)}_{n-k \text{ failures: } (1-p)^{n-k}}$$

$$= p^k (1-p)^{n-k}$$

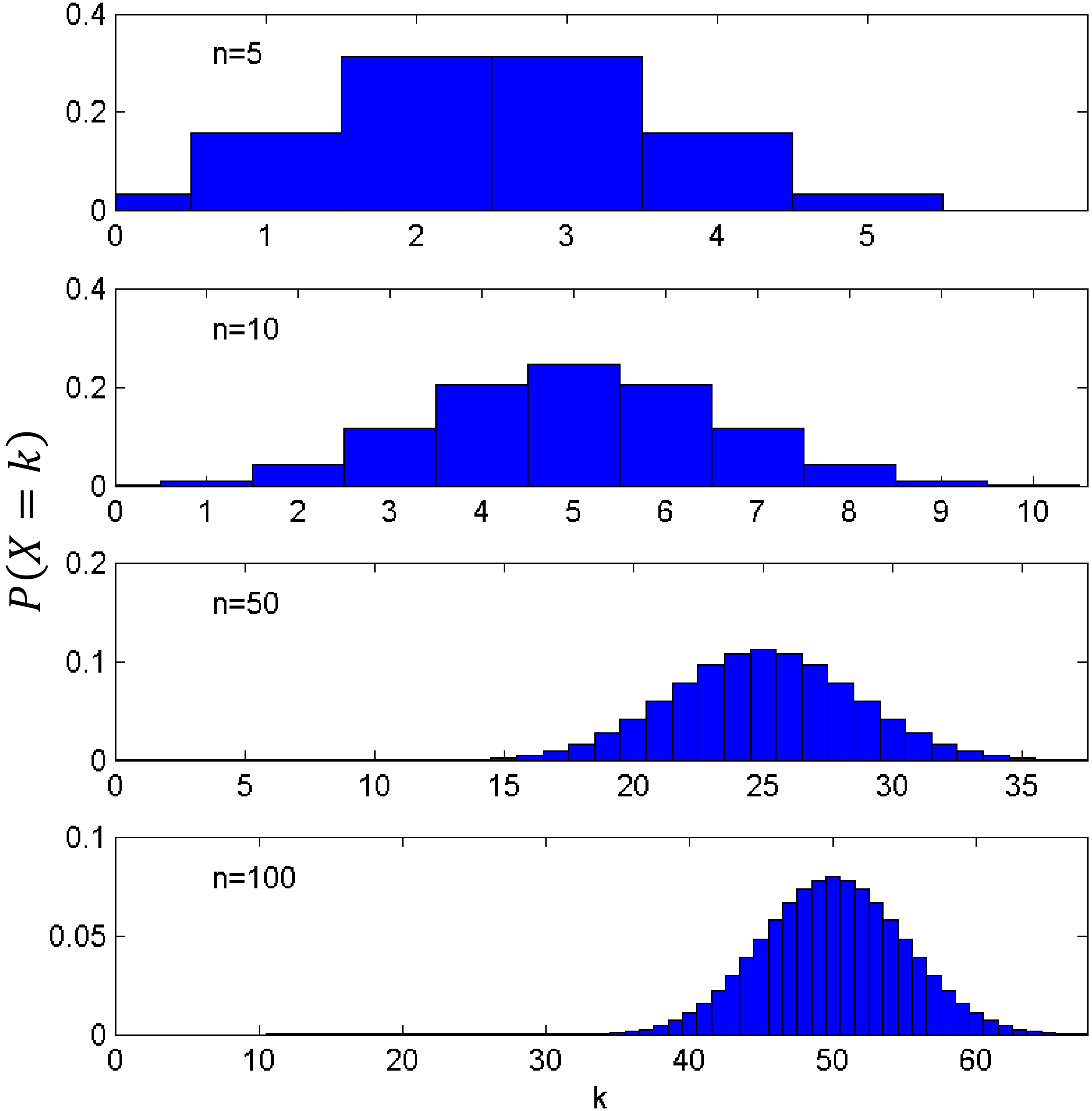
Every possible different ordering also has this same probability. The total number of ways of choosing *k* out of the *n* trails to be successes is $\binom{n}{k}$, so there are $\binom{n}{k}$, possible orderings.

Since each ordering is an exclusive possibility, by the special addition rule the overall probability is $p^k(1-p)^{n-k}$ added $\binom{n}{k}$ times:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$p = 0.5$$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$



Example: *If I toss a coin 100 times, what is the probability of getting exactly 50 tails?*

Answer:

Let X = number tails in 100 tosses

Bernoulli trial: tail or head, $X \sim B(n, p) = B(100, 0.5)$

$$P(X = 50) = C_k^n p^k (1-p)^{n-k} = C_{50}^{100} 0.5^{50} (1-0.5)^{50}$$

$$\approx 0.0796$$

Example: *A component has a 20% chance of being a dud. If five are selected from a large batch, what is the probability that more than one is a dud?*

Answer:

Let X = number of duds in selection of 5

Bernoulli trial: dud or not dud, $X \sim B(5,0.2)$

P(More than one dud) $= P(X > 1) = 1 - P(X \leq 1) = 1 - \mathrm{P}(\mathrm{X} = 0) - \mathrm{P}(\mathrm{X} = 1)$

$= 1 - C_0^5 0.2^0 (1 - 0.2)^5 - C_1^5 0.2^1 (1 - 0.2)^4$

$= 1 - 1 \times 1 \times 0.8^5 - 5 \times 0.2 \times 0.8^4$

= 1 - 0.32768 - 0.4096 ≈ 0.263.