Lecture 4 Probability

From set-based probability to random variables

10/6/2021

Lecture 3: Two by two tables

- Association between two binary variables
  - Relative Risk (相對風險)
  - Odds ratio (勝算比)

Example: 研究所入學管道與性別的關係 交大電 X 所 (七個學年度)

	女	男	Total
甄試	256	879	1135
考試	43	708	751
Total	299	1587	1886

Pr(女生) = 0.1585, Pr(男生) = 0.8415 → 男多女少的研究所

Pr(甄試) = 0.60, Pr(考試) = 0.40 → 甄試名額較考試多

Pr(甄試|女) = 0.8562, Pr(考試|女) = 0.1438 → 該所女生有八成五透過甄試

Pr(甄試|男) = 0.5539, Pr(考試|男) = 0.4461→ 該所男生有五成五透過甄試

Odds (甄試:考試) for 全體= $\frac{1135}{751} = 1.51$(全體學生:"甄試"是 "考試"人數的 1.51 倍)

Odds (甄試:考試) for 女生= $\frac{256}{43} = 5.95$(女生:"甄試"是 "考試"人數的 5.95 倍)

Odds (甄試:考試) for 男生= $\frac{879}{708} = 1.24$(男生:"甄試"是 "考試"人數的 1.24 倍)

Odds ratio (女:男) $= \frac{5.95}{1.24} = 4.7953$

解讀:

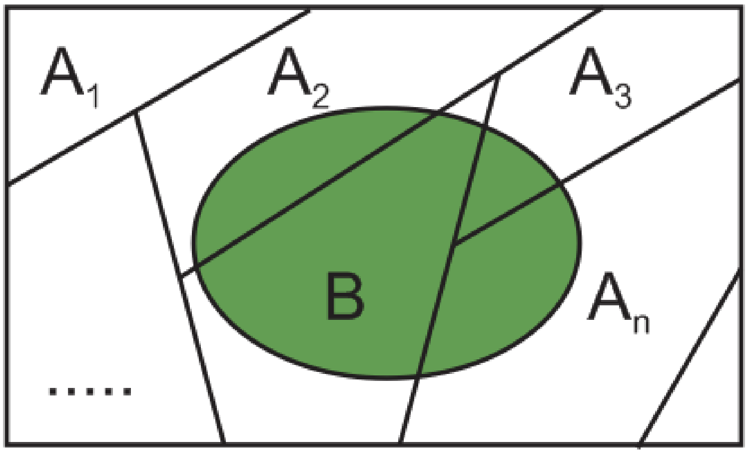
- Odds ratio 是否等於 1 是判斷相關性的關鍵 (不等於 1 代表有相關)

- 女生較男生容易透過甄試管道入學

補充: 多數交大研究所都有 Odds ratio > 1 的現象(已入學女生較高比例透過甄試)

猜測: 甄試重視大學 GPA

Bayes Theorem



Examples:

In a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05.
Suppose the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06.
What is the probability that an adult over 40 years of age is diagnosed as having cancer?

Let A: having cancer A^C: without cancer → partition sets

B: diagnosed of having cancer

$\Pr(A) = 0.05$ → $\Pr(A^C) = 0.95$

$\Pr(B \mid A) = 0.78$, $\Pr(B \mid A^C) = 0.06$ (題意似乎是這樣)

$$\begin{aligned}\Pr(\mathrm{B}) &= \Pr(B \mid A)\Pr(\mathrm{A}) + \Pr(B \mid A^C)\Pr(\mathrm{A}^C) \\ &= 0.78 \times 0.05 + 0.06 \times 0.95 \\ &= 0.096\end{aligned}$$

$$\Pr(\mathrm{A} \mid \mathrm{B}) = \frac{\Pr(B \mid A)\Pr(\mathrm{A})}{\Pr(B)} = \frac{0.78 \times 0.05}{0.096} = 0.40625$$

$$\Pr(\mathrm{A}^C \mid \mathrm{B}) = \frac{\Pr(B \mid A^C)\Pr(\mathrm{A}^C)}{\Pr(B)} = \frac{0.06 \times 0.95}{0.096} = 0.59375$$

Example

In a certain city, 30% of the people are conservatives, 50% are liberals,
and 20% are independents.
Records show that in a particular election, 65% of the conservatives voted,
82% of the liberals voted, and 50% of the independents voted.
If a person in the city is selected at random and *it is learned that she did not vote in the last election*, what is the probability that she is a Liberal?

$\Pr(A_1) =$ P(Conservatives) = 0.3;

$\Pr(A_2) =$ Pr(Liberals) = 0.5;

$\Pr(A_3) =$ Pr(independent) = 0.2.

B = "vote"

$\Pr(B \mid A_1) =$ 0.65; $\Pr(B \mid A_2) =$ 0.82; $\Pr(B \mid A_3) =$ 0.5.

Question: $\Pr(A_2 \mid B^C) =$ Pr(liberal | not vote) = ?

Hint: (計算不確定是否正確，請自行檢查)

1. $\Pr(B \mid A_1)\Pr(A_1) = \Pr(B \cap A_1)$ = 0.65* 0.3= 0.195

   $\Pr(B \mid A_2)\Pr(A_2) = \Pr(B \cap A_2)$ = 0.82*0.5= 0.41

   $\Pr(B \mid A_3)\Pr(A_3) = \Pr(B \cap A_3)$ =0.5*0.2=0.1

2. $\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \Pr(B \cap A_3)$ = 0.705

3. $\Pr(B^C) = 1 - \Pr(B)$ = 0.295

4. $\Pr(A_2 \mid B^C) = \dfrac{\Pr(A_2 \cap B^C)}{\Pr(B^C)} = \dfrac{\Pr(A_2) - \Pr(A_2 \cap B)}{1 - \Pr(B)} = \dfrac{0.5 - 0.41}{0.295}$ =0.3051

The obscure maths theorem that governs the reliability of Covid testing

The Observer
Coronavirus

Tom Chivers

Sun 18 Apr 2021 07.00 BST

https://www.theguardian.com/world/2021/apr/18/obscure-maths-bayes-theorem-reliability-covid-lateral-flow-tests-probability

貝氏定理與法律

Random variable (隨機變數)

Connection between set-based probability and random variables

➔ 從樣本空間找到可以量化 (測量或是分類) 的對應關係

Example: toss a coin three times

Two possible outcomes each time: H – (head, 正面) and T (tail, 反面)

Sample space: 8 elements $S = \{\omega_1, \ldots \omega_8\}$

$\omega_1 : HHH$, $\omega_2 : HHT$, $\omega_3 : HTH$, $\omega_4 : HTT$,

$\omega_5 : THH$, $\omega_6 : THT$, $\omega_7 : TTH$, $\omega_8 : TTT$

Define X = <u>number of heads obtained</u> in the three tosses

Possible values of X = 0, 1, 2, 3

$\{X = 0\} = \{(TTT)\}$.

$\{X = 1\} = \{(HTT), (THT), (TTH)\}$.

$\{X = 2\} = \{(HHT), (THH), (HTH)\}$.

$\{X = 3\} = \{(HHH)\}$.

Remarks:

- The value of X depends on the outcome of a random experiment.
- Thus we call X as a random variable.
- Formally, a random variable is a mapping from the sample space (the domain, 定義域) to the real line (range, 值域)

Example: $X(\omega_1) = 3$, $X(\omega_2) = 2$, $X(\omega_3) = 2$, $X(\omega_4) = 1$

$X(\omega_5) = 2$, $X(\omega_6) = 1$, $X(\omega_7) = 1$, $X(\omega_8) = 0$.

- Set-based probability focuses on the <u>domain</u> (sample space)

  while random-variable based probability focuses on the <u>range</u>.

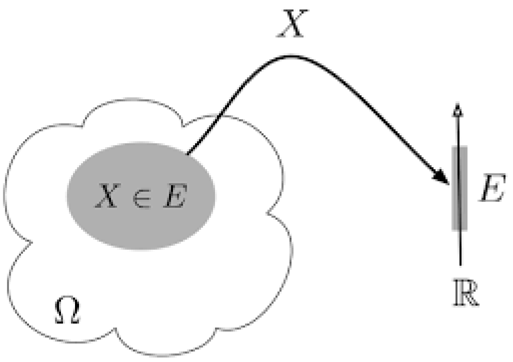
這個角度解釋了為何機率可以用集合討論 (以<u>定義域</u>的觀點), 或是以隨機變數來討論 (以<u>值域</u>的觀點).

Example: Pr(at least two heads) = $P(\{\omega_1, \omega_2, \omega_3, \omega_5\}) = \frac{4}{8}$

$$= P(X = 2) + P(X = 3) = \frac{3}{8} + \frac{1}{8}$$

Remarks:

- Probability based on random variables provides a simplified and systematic framework to study a phenomenon of interest.
- We can apply powerful mathematical skills to analyze real-world problems.



Types of random variables

1. discrete type (離散型)

   X = number of heads obtained in three tosses,

   X = number of babies born in 2012 in Taiwan

2. continuous type (連續型)

   X = height, blood pressure, temperature

3. mixed type (混合)

Definitions for discrete random variables

1. *The probability mass function (機率質量函數):*

$$p(x) = \Pr(X = x)$$

property: $\sum_{\forall x} \Pr(X = x) = 1$ (i.e. $\Pr(S) = 1$)

2. *The cumulative distribution of* X *(累積分佈函數):*

$$F(x) = \Pr(X \le x) = \sum_{a \le x} \Pr(X = a)$$

Properties:

- $F(x)$ is an increasing or non-decreasing function which jumps at "mass" point x with $\Pr(X = x) > 0$.

Remarks:

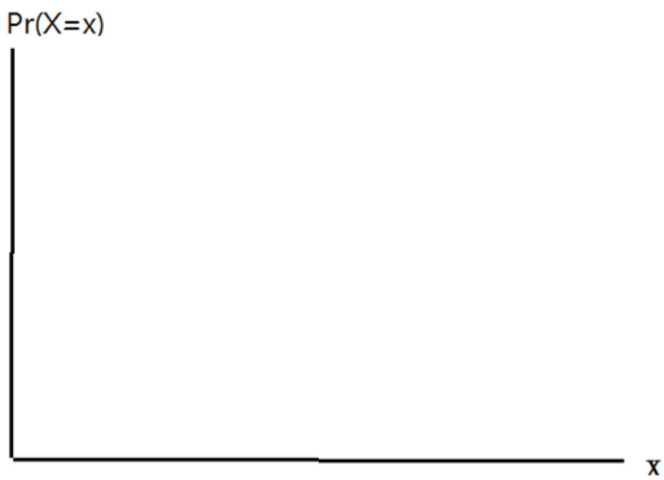
- $F(x)$ and $p(x)$ contain complete information of the random variable.
- 兩者可以互推，知道一個就可以推得另一個

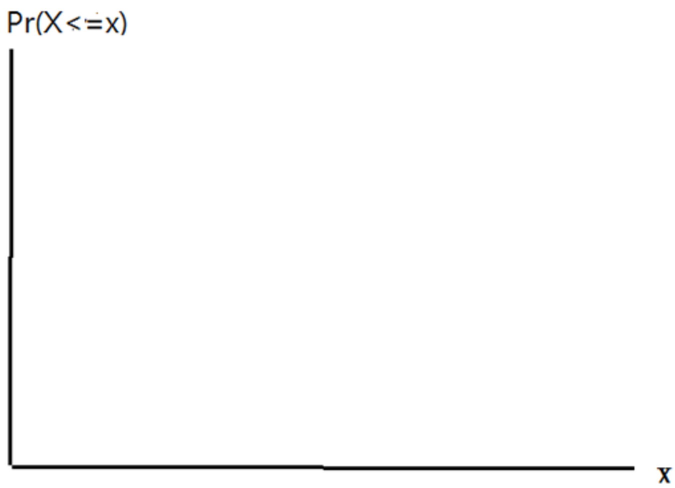
Example: Toss a coin three times, X = # of heads

$$\Pr(X = 0) = \frac{1}{8},\ \Pr(X = 1) = \frac{3}{8},\ \Pr(X = 2) = \frac{3}{8},\ \Pr(X = 3) = \frac{1}{8}$$

x value	Prob Pr(X=x)	Cumulative probability Pr(X <= x)
0	1/8	1/8
1	3/8	4/8
2	3/8	7/8
3	1/8	1

Draw Pr(X = x) and $F(x)$ (probability and distribution functions)





Exercise:

The following table shows the density for the random variable X, the number of wing beats per second of a species of large moth while in flight.

x	6	7	8	9	10
$f(x)$	.05	.1	.6	.15	?

(*a*) Find $f(10)$.
(*b*) Find $P[X \leq 8]$. Interpret this probability in the context of this problem.
(*c*) Find $P[X < 8]$.
(*d*) Find $P[X \geq 7]$.
(*e*) Find $P[X > 7]$.

Conventions of notations: (統計領域的符號使用習慣，目的是便於快速理解)

- Capital English letter (大寫英文字母): random variables
- Lower-case English letter (小寫英文字母):
  - realizations of random variables which are real numbers
- Greek letter (希臘字):
  - parameters which are constants
  - In probability calculations, parameter values are given; while in reality, parameters are unknown whose values can be estimated using statistical methods.

機率的計算會假設參數值“已知”，統計會將參數視為“未知”，予以估計

Useful parameters of a (discrete) random variable

1. Expected value of X = the mean of X → information about the "center"

$$\mu = E(X) = \sum_x x \Pr(X = x): \text{“期望值” 或是 “平均數”}$$

2. The variance of X (變異數) → information about the "dispersion"

$$\sigma^2 = Var(X) = \sum_x (x-\mu)^2 \Pr(X = x)$$

$$= E(X^2) - \mu^2$$ → easier to work with

Proof: $Var(X) = \sum_x (x-\mu)^2 \Pr(X = x) = \sum_x \{x^2 - 2\mu x + u^2\} \Pr(X = x)$

$$= \left[\sum_x x^2 \Pr(X = x)\right] - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$$

3. The rth moment of X (X 的 r 階動差)

$$E(X^r) = \sum_x x^r \Pr(X = x)$$

$E(X) = \mu$ → the first moment is the mean

$E(X^2) = \mu^2 + \sigma^2$

→ the second moment contains the information of mean and variance

$E(X^3)$ → the third moment is related to skewness

$E(X^4)$ → the third moment is related to kurtosis (尾巴的厚度)

4. $E[g(X)]$ (X 函數的期望值)

$g(.)$ is a function

$$E[g(X)] = \sum_{\forall x} g(x) \Pr(X = x)$$

Example 1: $g(X) = X$: $\mu = E(X) = \sum_{\forall x} x \Pr(X = x)$ → mean

Example 2: $g(X) = (X - \mu)^2$

■ $\sigma^2 = Var(X) = E[(X - \mu)^2] = \sum_{\forall x} (x - \mu)^2 \Pr(X = x)$

$= E(X^2) - \mu^2$

Remark: Properties of expectation

X and Y are random variables, a, b, c are constants

- $E(a) = a$ and $Var(a) = 0$
- $E(aX + bY) = aE(X) + bE(Y)$
- $Var(aX) = a^2 Var(X)$
- Covariance → later

Example 3: $g(X) = e^{tX}$ (**Moment generating function, 動差母函數)**

■ $E[e^{tX}] = \sum_{\forall x} e^{tx} \Pr(X = x) = M_X(t)$ $(t \in R)$ → function of t

■ We can derive the moments of X from $M_X(t)$

$\frac{\partial M_X(t)}{\partial t}\Big|_{t=0} = E(X)$ → 一次微分, t 代 0, 得一階動差

$\frac{\partial^2 M_X(t)}{\partial t^2}\Big|_{t=0} = E(X^2)$ → 二次微分, t 代 0, 得二階動差

…

■ It is easier to take derivatives of a function that compute the moments directly. The former requires compute the sums or perform integrations; while the latter involves taking derivatives (微分比積分簡單)

Remark:

Calculations of "Average" of the sample and "Mean" of a discrete random variables

Note: 為何 資料分析裡的平均數與"機率"裡的 mean 計算公式乍看不同?

- In the sample (data), 所有資料點出現的可能性都一樣, 所以機率是 1/n.
- Under the framework of probability, $\Pr(X = x)$ is assumed to be known.

Example: $X =$ **the score of Essay Writing for the entrance exam (會考作文)**

= 0,1,…,6

Made-up example (Suppose there are 100 students taking the exam)

score	**0**	**1**	**2**	**3**	**4**	**5**	**6**	Total
number	0	15	18	25	24	16	2	100
probability	0	.15	.18	.25	.24	.16	0.02	1

Sample average

$$\bar{X} = \frac{\sum_{i=1}^{100} x_i}{100} = \frac{0*0+1*15+2*18+3*25+4*24+5*16+6*2}{100}$$

$$= \sum_{x=0}^{6} x\Pr(X = x) = 0*0+1*0.15+2*018+3*0.25+4*0.24+5*0.16+6*0.02$$

News (實際的數字): 105 年寫作測驗考「從陌生到熟悉」

6 級分人數共 4972 人（去年 3585 人），占整體 1.8%、

5 級分有 52563 人，占 19.08%、

4 級分人數最多有 173542 人，占 62.99%、

3 級分有 27112，占 9.84%、

2 級分有 7214 人，占 2.62%、

1 級分 2611 人，占 0.95%、

0 級分 7501 人 2.72%。

- **Special discrete variables (因為有用，所以會予以命名)**

- They are useful to describe the real world phenomenon
- Thus their properties are thoroughly discussed
- Given their names, you know what they are.
- 對學生來說，最困難的是遇到實際問題，不知道使用哪個分配去描述

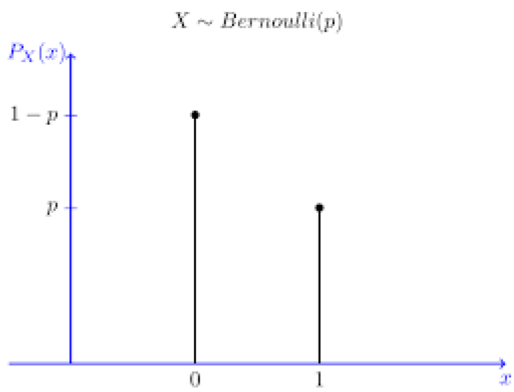
1. Bernoulli random variable (白努力隨機變數)

Bernoulli trial: a trial which contains two possible outcomes (success or failure)

Define

$Y = 1$ if "success" occurs

$= 0$ if "failure" occurs



Probability function of $Y \sim \text{Bernoulli}(p)$

- $\Pr(Y=1) = p$ and $\Pr(Y=0) = 1-p$
- Write the above two formula in one equation:

$$\Pr(Y = y) = p^y (1-p)^{1-y} \quad \text{for} \quad y = 0,1$$

Derivations:

$$E(Y) = 1 \times \Pr(Y=1) + 0 \times \Pr(Y=1) = p$$

$$E(Y^2) = 1^2 \times \Pr(Y=1) + 0^2 \times \Pr(Y=1) = p$$

$$Var(Y) = E(Y^2) - [E(Y)]^2 = p - p^2 = p(1-p)$$

Remarks:

- $Var(Y)$ achieves its maximum if $p = 1-p = \dfrac{1}{2}$

  (旗鼓相當的比賽，結果最難預料)

- Odds = $\dfrac{p}{1-p}$

2. Binomial random variable (二項式隨機變數)

滿足 Binomial 分配的基本要素 (Binomial Setting)

THE BINOMIAL SETTING

1. There are a fixed number n of observations.
2. The n observations are all **independent.** That is, knowing the result of one observation tells you nothing about the other observations.
3. Each observation falls into one of just two categories, which for convenience we call "success" and "failure."
4. The probability of a success, call it p, is the same for each observation.

Determine whether the random variable defined in each example is a Binomial random variable.

Example: A couple have 3 children, X = # of boys

Ans: YES

Example: A couple decides to continue to have children until their first girl is born.
X = # of children

Ans: NO, n is not fixed.

Example: Randomly select 10 persons from a sample consisting of 20 males and 20 females. X = # of males

Ans: NO, the trials are not independent; the success probability is different in each selection. (每次機率不同，與前面結果有關)

Example: Randomly select 10 persons from a sample consisting of 5000 males and 5000 females. X = # of males

Ans: X "approximately" follows a Binomial distribution . (每次機率不同，與前面結果有關，但因為總數很大，影響很小)

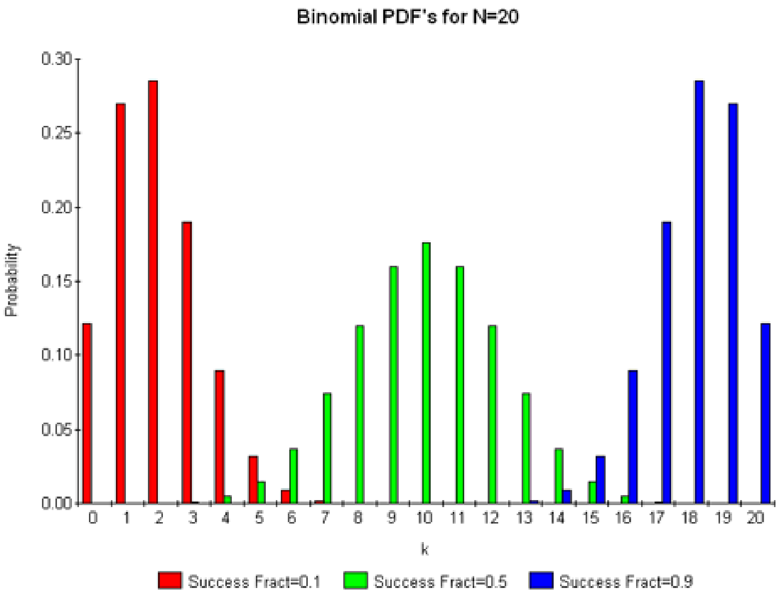
Probability function for $X \sim Binomial(n, p)$:

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for} \quad x = 0,1,\ldots,n$$

Derivations:

- the probability of getting $(S, S, \ldots, S, F, F, \ldots, F)$ → $p^x(1-p)^{n-x}$
- the probability of getting $(S, F, \ldots, S, S, F, \ldots, F)$ → $p^x(1-p)^{n-x}$
- In n trials, the number of combinations for x successes and $(n-x)$ failure → $\binom{n}{x}$

- probability of the union of the $\binom{n}{x}$ disjoint events = the sum of each probability ($p^x(1-p)^{n-x}$)
- Left: $X \sim Binomial(n = 20, p = 0.1)$
- Middle: $X \sim Binomial(n = 20, p = 0.5)$ → symmetric
- Right: $X \sim Binomial(n = 20, p = 0.9)$



C. Important properties of X ~ Binomial(n,p)

$$E(X)=\sum_{x=0}^{x=n} x\cdot\binom{n}{x}p^x(1-p)^{n-x}=n\cdot p$$

$$Var(X)=\sum_{x=0}^{x=n} x^2\cdot\binom{n}{x}p^x(1-p)^{n-x}-(n\cdot p)^2=n\cdot p\cdot(1-p)$$

$$F(a)=\sum_{x=0}^{x=a} x\cdot\binom{n}{x}p^x(1-p)^{n-x} \quad \text{(no explicit form)}$$

skill: $\Pr(X\geq 1)=1-\Pr(X=0)$ (右式比較好算)

Proof: (2 ways)

a. direct proof → not easy to calculate the sum of a complicated series

(數學上求級數和，並不容易，往往沒有 closed-form)

b. From its relationship with Bernoulli random variables → easy

令 $X=Y_1+\ldots+Y_n$, where $Y_i \sim Bernoulli(p)$.

know: $E(Y_i)=1\cdot p+0\cdot(1-p)=p$

$$Var(Y_i)=1^2\cdot p+0^2\cdot(1-p)-p^2=p(1-p)$$

→ $E(X)=E(Y_1+\ldots+Y_n)=n\cdot p$

→ $Var(X)=Var(Y_1+\ldots+Y_n)=\sum_{i=1}^{n}Var(Y_i)+\sum\sum_{i\neq j}Cov(Y_i,Y_j)$

$=n\cdot p(1-p)+0$ (note Y_i & Y_j are independent, second term = 0)

Useful properties for expectations (revisited)

1.

$E(X+Y)=E(X)+E(Y)$

$E(c)=c$ (常數的期望值)

2. When c is a constant, $Var(c)=0$

Supplementary Knowledge: Two random variables (will not be on the exam)

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_Y)]$$

$$= \sum_x \sum_y (x-\mu_x)(y-\mu_y)\Pr(X=x, Y=y)$$

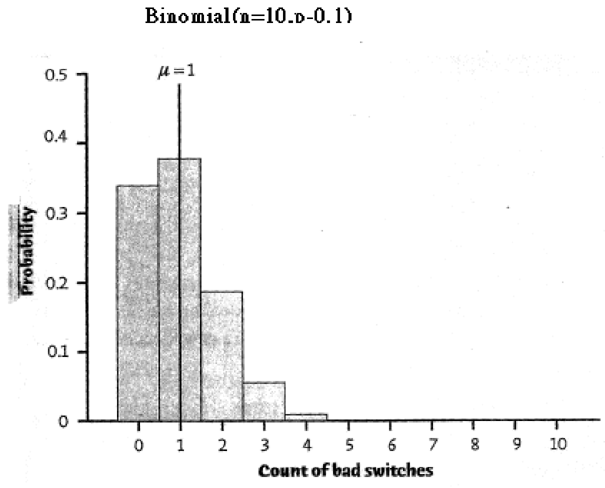
Correlation coefficient:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E[(X-\mu_x)(Y-\mu_Y)]}{\sqrt{E[(X-\mu_x)^2]E[(Y-\mu_Y)^2]}}$$

Remarks:

- If X and Y are independent → $Cov(X,Y)=0$ but not vice versa.
- $Var(X) = Cov(X,X)$

Exercises:



Example: Inspecting switches

Continuing Example 12.5, the count X of bad switches is binomial with $n = 10$ and $p = 0.1$. The histogram in Figure 12.2 displays this probability distribution. (Because probabilities are long-run proportions, using probabilities as the heights of the bars shows what the distribution of X would be in very many repetitions.) The distribution is strongly skewed. Although X can take any whole-number value from 0 to 10, the probabilities of values larger than 5 are so small that they do not appear in the histogram.

The mean and standard deviation of the binomial distribution in Figure 12.2 are

$$\mu = np = (10)(0.1) = 1$$

$$\sigma = \sqrt{np(1-p)}$$

$$= \sqrt{(10)(0.1)(0.9)} = \sqrt{0.9} \doteq 0.9487$$

XAMPLE 12.1 Blood types

Genetics says that children receive genes from their parents independently. Each child of a particular pair of parents has probability 0.25 of having type O blood. If these parents have 5 children, the number who have type O blood is the count X of successes in 5 independent trials with probability 0.25 of a success on each trial. So X has the binomial distribution with $n = 5$ and $p = 0.25$.

Pr(<u>None</u> of the children have blood type O) $= \left(\frac{3}{4}\right)^5$

Pr(<u>Exactly</u> one child has blood type O) = $\binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4$

Pr(<u>At least</u> two children have blood type O) = $1 - \mathrm{P}(X = 0) - P(X = 1)$

$\mathrm{P}(X = 0) = \binom{5}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^5$; $\mathrm{P}(X = 1) = \binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4$

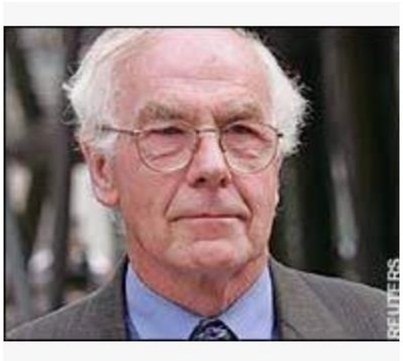
Pr(<u>At most</u> two children have blood type O) = $\mathrm{P}(X = 0) + P(X = 1) + P(X = 2)$

$\mathrm{P}(X = 2) = \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3$

法律裡的機率

Sally Clark (維基百科) 1964 – 2007

- Sally Clark was an English solicitor who, in November 1999, became the victim of a miscarriage of justice when she was found guilty of the murder of her two infant sons. Clark's first son died in December 1996 within a few weeks of his birth, and her second son died in similar circumstances in January 1998. A month later, Clark was arrested and tried for both deaths.
    - ■ The defence argued that the children had died of sudden infant death syndrome (SIDS). 嬰兒猝死症
    - ■ The prosecution case relied on flawed statistical evidence presented by paediatrician Professor Sir Roy Meadow, who testified that the chance of two children from an affluent family suffering SIDS was 1 in 73 million. He had arrived at this figure by squaring his estimate of a chance of 1 in 8500 of an individual SIDS death in similar circumstances.



Meadow did mislead the Sally Cl...
telegraph.co.uk

    - ■ The Royal Statistical Society later issued a statement arguing that there was no statistical basis for Meadow's claim, and expressed concern at the "misuse of statistics in the courts".

    - ■ Clark was convicted in November 1999. m
- The convictions were upheld on appeal in October 2000, but overturned in a second appeal in January 2003, after it emerged that Alan Williams, the prosecution forensic pathologist who examined both babies, had failed to disclose microbiological reports that suggested the second of her sons had died of natural causes.
    - ■ Clark was released from prison having served more than three years of her sentence. Journalist Geoffrey Wansell called Clark's experience "one of the great miscarriages of justice in modern British legal history".
    - ■ As a result of her case, the Attorney General Lord Goldsmith ordered a review of hundreds of other cases, and two other women had their convictions overturned.

- Clark's experience caused her to develop severe psychiatric problems and she died in her home in March 2007 from alcohol poisoning.

Statistical evidence

- The first trial was widely criticised for the misrepresentation of statistical evidence, particularly by Meadow.
  - He stated in evidence as an expert witness that "one sudden infant death in a family is a tragedy, two is suspicious and three is murder unless proven otherwise" (Meadow's law).
    - He claimed that, for an affluent non-smoking family like the Clarks, the probability of a single cot death was 1 in 8,543, **so the probability of two in the same family was around "1 in 73 million" (8543 × 8543).**
    - Given that there are around 700,000 live births in Britain each year, Meadow argued that **a double cot death would be expected to occur once every hundred years.**

- In October 2001, the Royal Statistical Society (RSS, 英國皇家統計協會) issued a public statement expressing its concern at the "misuse of statistics in the courts".
  - It noted that there was "no statistical basis" for the "1 in 73 million" figure. In January 2002, **the RSS wrote to the Lord Chancellor pointing out that "the calculation leading to 1 in 73 million is false"**.
    - **Meadow's calculation was based on the assumption that two SIDS deaths in the same family are independent. The RSS argued that "there are very strong reasons for supposing that the assumption is false.** There may well be unknown genetic or environmental factors that predispose families to SIDS, so that a second case within the family becomes much more likely than would be a case in another, apparently similar, family." The prosecution did not provide any evidence to support its different assumption.

- In a 2004 article in Paediatric and Perinatal Epidemiology, Professor of Mathematics Ray Hill of Salford University concluded, using extensive SIDS statistics for England, that **"after a first cot death the chances of a second become greatly increased" by a dependency factor of between 5 and 10**.
- The ruling was also subject to a statistical error known as the "prosecutor's fallacy".

檢察官謬誤（**prosecutor's fallacy**）是一種非形式謬誤，係取一不甚相關、或有關但未正確考慮條件機率的數據，認定被告「無辜的機率」很小。

- Many press reports of the trial reported that the "1 in 73 million" figure was the probability that Clark was innocent. However, even if the "1 in 73 million" figure were valid, this should not have been interpreted as the probability of Clark's innocence.
- In order to calculate the probability of Clark's innocence, the jury needed to weigh up the relative likelihood of the two competing explanations for the children's deaths. In other words, murder was not the only alternative explanation for the deaths in the way that one might have inferred from asserted probability of double SIDS.
- Although double SIDS is very rare, double infant murder is likely to be rarer still, so the probability of Clark's innocence was quite high. **Hill calculated the odds ratio for double SIDS to double homicide at between 4.5:1 and 9:1.**

- Hill raised a third objection to the "1 in 73 million" figure: the probability of a child dying from SIDS is 1 in 1,300, not 1 in 8,500.[clarification needed] Meadow arrived at the 1 in 8,500 figure by taking into account three key characteristics possessed by the Clark family, all of which make SIDS less likely.
- However, Hill said that Meadow "conveniently ignored factors such as both the Clark babies being boys – which make cot death more likely".

- During the second appeal**, the court noted that Meadow's calculations were subject to a number of qualifications, but "none of these qualifications were referred to by Professor Meadow in his evidence to the jury and thus it was the headline figures of 1 in 73 million that would be uppermost in the jury's minds". (背後的假設沒說清楚)**
- The appeal court concluded that "the evidence should never have been before the jury in the way that it was when they considered their verdicts". The judges continued, "we rather suspect that with the graphic reference by Professor Meadow to the chances of backing long odds winners of the Grand National year after year it may have had a major effect on [the jury's] thinking notwithstanding the efforts of the trial judge to down play it".

【利物浦】Grand National 全國來賭馬！

Despite feeling stressed out with course work, this was a chance I was not ... 因為本日是『Grand National』，全國平時不賭馬的都會選再這天碰碰運氣，所以人潮 ...