Conditional Probability Calculations and Joint Probability Distributions

Parthanil Roy

Indian Statistical Institute

March 24, 2017

What is a "Conditional Probability"?

What is a "Conditional Probability"?

- If A and B are two events with $P(B) > 0$ then the conditional probability of A given B is denoted by $P(A|B)$ and defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

If $P(B) = 0$, then the conditional probability of A given B is *not defined*.

What is a “Conditional Probability”?

- If A and B are two events with $P(B) > 0$ then the conditional probability of A given B is denoted by $P(A|B)$ and defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

If $P(B) = 0$, then the conditional probability of A given B is *not defined*.

- $P(A|B)$ measures the chance that A will occur if it is known that B has already occurred.

The Chain Rule

Multiplication Rule. Directly from the definition of conditional probability, it follows that

$$P(A \cap B) = P(B)P(A|B)\,.$$

The multiplication rule has a very useful extension to more than 2 events that is also called *the chain rule*.

Chain Rule. Let A_1, A_2, A_3 be arbitrary events such that $P(A_1 \cap A_2) > 0$. Then

$$\begin{aligned} &P(A_1 \cap A_2 \cap A_3) \\ &\quad = P(A_1 \cap A_2)P(A_3|A_1 \cap A_2) \\ &\quad = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \end{aligned}$$

The First Example

The First Example

Example

Two mountain resorts S and T are connected by a direct road, and a more circuitous route going through a third resort U. Suppose A_1, A_2 and A_3 are the events that roads ST, SU and TU are open on a typical winter day, and

$$P(A_1) = .4, \quad P(A_2) = .75, P(A_3|A_2) = .8, \quad P(A_1|A_2 \cap A_3) = .5.$$

What is the probability that on a typical winter day a traveller will be able to get from S to T?

The First Example

Example

Two mountain resorts S and T are connected by a direct road, and a more circuitous route going through a third resort U. Suppose A_1, A_2 and A_3 are the events that roads ST, SU and TU are open on a typical winter day, and

$$P(A_1) = .4, \quad P(A_2) = .75, P(A_3|A_2) = .8, \quad P(A_1|A_2 \cap A_3) = .5.$$

What is the probability that on a typical winter day a traveller will be able to get from S to T?

Solution. We are looking for $P(A_1 \cup (A_2 \cap A_3))$.

The First Example

Example

Two mountain resorts S and T are connected by a direct road, and a more circuitous route going through a third resort U. Suppose A_1, A_2 and A_3 are the events that roads ST, SU and TU are open on a typical winter day, and

$$P(A_1) = .4, \quad P(A_2) = .75, P(A_3|A_2) = .8, \quad P(A_1|A_2 \cap A_3) = .5.$$

What is the probability that on a typical winter day a traveller will be able to get from S to T?

Solution. We are looking for $P(A_1 \cup (A_2 \cap A_3))$. By inclusion-exclusion

$$P(A_1 \cup (A_2 \cap A_3)) = P(A_1) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3).$$

The First Example

Example

Two mountain resorts S and T are connected by a direct road, and a more circuitous route going through a third resort U. Suppose A_1, A_2 and A_3 are the events that roads ST, SU and TU are open on a typical winter day, and

$$P(A_1) = .4, \quad P(A_2) = .75, P(A_3|A_2) = .8, \quad P(A_1|A_2 \cap A_3) = .5.$$

What is the probability that on a typical winter day a traveller will be able to get from S to T?

Solution. We are looking for $P(A_1 \cup (A_2 \cap A_3))$. By inclusion-exclusion

$$P(A_1 \cup (A_2 \cap A_3)) = P(A_1) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3).$$

By the chain rule

$$P(A_2 \cap A_3) = P(A_2)P(A_3|A_2) = .75 \times .8 = .6,$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_2)P(A_3|A_2)P(A_1|A_2 \cap A_3) = .75 \times .8 \times .5 = .3.$$

The First Example

Example

Two mountain resorts S and T are connected by a direct road, and a more circuitous route going through a third resort U. Suppose A_1, A_2 and A_3 are the events that roads ST, SU and TU are open on a typical winter day, and

$$P(A_1) = .4, \quad P(A_2) = .75, P(A_3|A_2) = .8, \quad P(A_1|A_2 \cap A_3) = .5.$$

What is the probability that on a typical winter day a traveller will be able to get from S to T?

Solution. We are looking for $P(A_1 \cup (A_2 \cap A_3))$. By inclusion-exclusion

$$P(A_1 \cup (A_2 \cap A_3)) = P(A_1) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3).$$

By the chain rule

$$P(A_2 \cap A_3) = P(A_2)P(A_3|A_2) = .75 \times .8 = .6,$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_2)P(A_3|A_2)P(A_1|A_2 \cap A_3) = .75 \times .8 \times .5 = .3.$$

Therefore, the required probability is $.4 + .6 - .3 = .7$.

The Law of Total Probability and Bayes' Formula

Let $A_1, A_2, \ldots, ...A_n$ be a *partition of the sample space* Ω. This means that

1. $A_1, A_2, \ldots, ...A_n$ are pairwise disjoint, and
2. $\cup_{j=1}^n A_j = \Omega$.

Then the following hold for any event B:

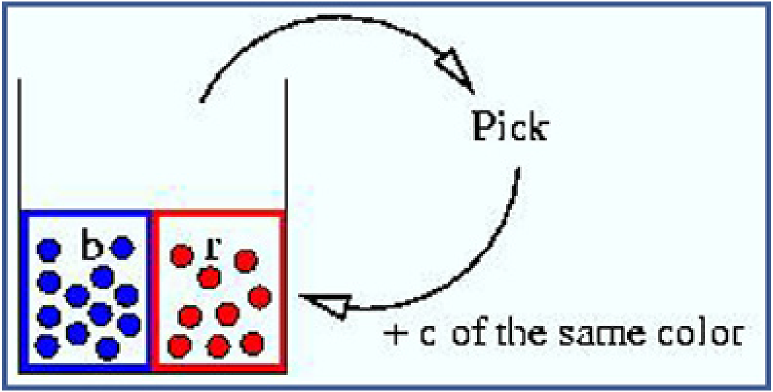
1. **Law of Total Probability.**

$$P(B) = \sum_{j=1}^n P(B \cap A_j) = \sum_{j=1}^n P(A_j) P\left(B \middle| A_j\right).$$

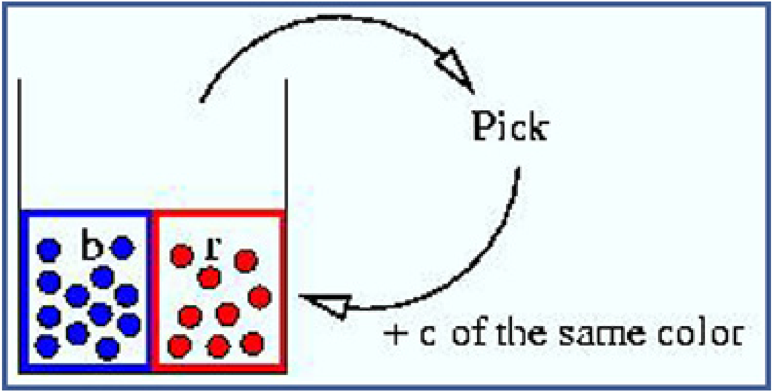
2. **Bayes Formula.**

$$P\left(A_i \middle| B\right) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) P\left(B \middle| A_i\right)}{\sum_{j=1}^n P(A_j) P\left(B \middle| A_j\right)}.$$

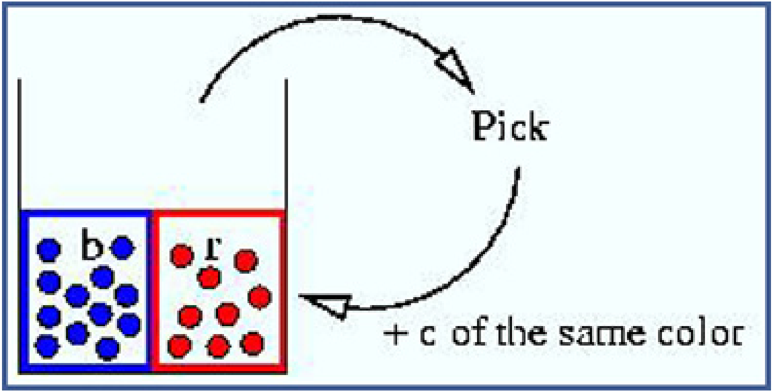
Example: Polya's Urn Scheme



P(2nd Ball is blue)=?



P(1st Ball was Red | 2nd Ball is blue)=?



The Betting Game

Suppose there are n people in a room and I make a bet claiming that there are at least two people having the same birthday. What is the probability that I shall win the bet?

To solve this problem, make the following assumptions:

- No birthday on 29th Feb.
- All birthdays are independent.
- All 365 days of a year are equally likely to be a birthday.

Solution to the Betting Game

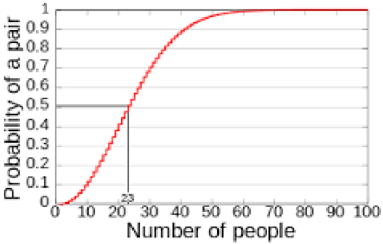
Suppose there are n people in a room and I make a bet claiming that there are at least two people having the same birthday. What is the probability that I shall win the bet?

Under the assumptions, the required probability

$$
\begin{aligned}
&= 1 - P(\text{all n people have different birthdays}) \\
&= 1 - \frac{\text{No. of ways n people will have different birthdays}}{\text{Total no. of ways}} \\
&= 1 - \frac{365 \times 364 \times 363 \times \cdots \times (365 - n + 1)}{(365)^n}
\end{aligned}
$$

Betting Game: The Graph

Suppose there are n people in a room and I make a bet claiming that there are at least two people having the same birthday. What is the probability that I shall win the bet? In other words, what is the probability that there is at least one pair in the room having the same birthday?



Betting Game: The Table

Suppose there are n people in a room and I make a bet claiming that there are at least two people having the same birthday. What is the probability that I shall win the bet? In other words, what is the probability (= p(n), say) that there is at least one pair in the room having the same birthday?

n	$p(n)$
1	0.0%
5	2.7%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
40	89.1%
50	97.0%
60	99.4%

Betting Game Involving Money

Suppose there are 50 people in a room and I make a bet claiming that there are at least two people having the same birthday. If I lose the bet, I shall give you Rs. 10000, and if you lose, you will only have to give me Rs. 1000. Will you play the game?

Betting Game Involving Money

Suppose there are 50 people in a room and I make a bet claiming that there are at least two people having the same birthday. If I lose the bet, I shall give you Rs. 10000, and if you lose, you will only have to give me Rs. 1000. Will you play the game?

n	*p*(*n*)
1	0.0%
5	2.7%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
40	89.1%
50	97.0%
60	99.4%

Betting Game Involving Money

Suppose there are 50 people in a room and I make a bet claiming that there are at least two people having the same birthday. If I lose the bet, I shall give you Rs. 20000, and if you lose, you will only have to give me Rs. 1000. Will you play the game?

n	*p*(*n*)
1	0.0%
5	2.7%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
40	89.1%
50	97.0%
60	99.4%

Betting Game Involving Money

Suppose there are 50 people in a room and I make a bet claiming that there are at least two people having the same birthday. If I lose the bet, I shall give you Rs. 30000, and if you lose, you will only have to give me Rs. 1000. Will you play the game?

n	$p(n)$
1	0.0%
5	2.7%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
40	89.1%
50	97.0%
60	99.4%

Expectation and Its Meaning

$$E(X) = \mu_X = \sum_x x P(X = x) = \sum (\text{value} \times \text{probability})$$

Expectation and Its Meaning

$$E(X) = \mu_X = \sum_x xP(X = x) = \sum (\text{value} \times \text{probability})$$

It measures the average / mean / expected value of a random variable.

Expectation and Its Meaning

$$E(X) = \mu_X = \sum_x x P(X = x) = \sum (\text{value} \times \text{probability})$$

It measures the average / mean / expected value of a random variable.

Just think of the betting game:

Expectation and Its Meaning

$$E(X) = \mu_X = \sum_x xP(X = x) = \sum (\text{value} \times \text{probability})$$

It measures the average / mean / expected value of a random variable.

Just think of the betting game: If X (in Rs.) is your profit, then $E(X)$ is your *expected profit / profit on the average.*

Discrete Random Vectors

If X and Y are discrete random variables, with possible values $\{x_i : i = 1, 2, \ldots, m\}$ and $\{y_j : j = 1, 2, \ldots, n\}$ respectively, then the random vector (X, Y) is called discrete.

Discrete Random Vectors

If X and Y are discrete random variables, with possible values $\{x_i : i = 1, 2, \ldots, m\}$ and $\{y_j : j = 1, 2, \ldots, n\}$ respectively, then the random vector (X, Y) is called discrete.

Joint probability mass function (**joint pmf**) of a discrete random vector (X, Y) (or, of jointly distributed discrete random variables X and Y) is

$$p_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j).$$

for all $i = 1, 2, \ldots, m$ and for all $j = 1, 2, \ldots, n$.

An Example of a Discrete Random Vector

In an automobile plant two tasks are performed by robots. The first task entails welding two joints, the second, tightening three bolts. Let X denote the number of defective welds, and Y the number of improperly tightened bolts produced per car. From the past history the joint pmf of X and Y is

x_i/y_j	0	1	2	3
0	.840	.030	.020	.010
1	.060	.010	.008	.002
2	.010	.005	.004	.001

Compute the probability that exactly 3 errors are made.

An Example of a Discrete Random Vector

In an automobile plant two tasks are performed by robots. The first task entails welding two joints, the second, tightening three bolts. Let X denote the number of defective welds, and Y the number of improperly tightened bolts produced per car. From the past history the joint pmf of X and Y is

x_i/y_j	0	1	2	3
0	.840	.030	.020	.010
1	.060	.010	.008	.002
2	.010	.005	.004	.001

Compute the probability that exactly 3 errors are made.

Answer:

$$P(X+Y=3) = P(X=0, Y=3)+$$

$$+P(X=1, Y=2) + P(X=2, Y=1)$$

$$= p_{X,Y}(0,3) + p_{X,Y}(1,2) + p_{X,Y}(2,1)$$

$$= .010 + 0.008 + 0.005 = .023.$$

Covariance and Its Meaning

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

Covariance and Its Meaning

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

It measures the amount (and direction) of liner association between X and Y.

Covariance and Its Meaning

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

It measures the amount (and direction) of liner association between X and Y.

However, it is not unit free:

Covariance and Its Meaning

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

It measures the amount (and direction) of liner association between X and Y.

However, it is not unit free: $Cov(aX, bY) = abCov(X, Y)$.

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

It measures the amount (and direction) of liner association between X and Y.

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

It measures the amount (and direction) of liner association between X and Y.

It is unit free:

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

It measures the amount (and direction) of liner association between X and Y.

It is unit free: $Corr(aX, bY) = Corr(X,Y)$.

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

It measures the amount (and direction) of liner association between X and Y.

It is unit free: $Corr(aX, bY) = Corr(X,Y)$.

It is always between -1 and +1. Both of these values are possible.

Correlation and Its Meaning

$$Corr(X,Y) = \rho(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

It measures the amount (and direction) of liner association between X and Y.

It is unit free: $Corr(aX, bY) = Corr(X,Y)$.

It is always between -1 and +1. Both of these values are possible. When?