RWTH BUSINESS SCHOOL

Mathematics & Statistics
M.Sc. Data Analytics and Decision Science
Prof. Dr. Thomas S. Lontzek

AACSB ACCREDITED

BUSINESS SCHOOL | RWTH AACHEN UNIVERSITY

The Prosecutor’s Fallacy

The ”Prosecutor's Fallacy” is a flaw in logical thinking: a confusion about what has happened and what is still uncertain.

In probability terms, it is a confusion of conditional probabilities $P(\mathrm{A}|\mathrm{B})$ and $P(\mathrm{B}|\mathrm{A})$.

That confusion has extreme implications in court. There are very many real-life occurrences of that fallacy with devastating results.

The Prosecutor's Fallacy - Example

Consider the following two possible outcomes:

- I = defendant is innocent
- DE = damning evidence, e.g., a DNA match at the crime scene.

Consider also the following conditional probabilities:

$P(\mathrm{DE}|\mathrm{I}) =$ probability that an innocent person matches the damning evidence

$P(\mathrm{I}|\mathrm{DE}) =$ probability that a person with the damning evidence is innocent

The Prosecutor's Fallacy - Example

The prosecution's expert witness states:

"An innocent person has a 1 in 100,000 chance of matching the damning evidence (DE)"

The prosecutor concludes:

- "The defendant has the damning evidence (DE)"
- "Therefore, there is a 1 in 100,000 chance that the defendant is innocent"
- "Clearly, the defendant is guilty"

The Prosecutor's Fallacy - Example

The prosecution’s expert witness states:

“An innocent person has a 1 in 100,000 chance of matching the damning evidence (DE)”

The prosecutor concludes:

- “The defendant has the dam[illegible] vidence (DE)”
- “Therefore, there is a 1 [illegible] 000 chance that the defendant is innocent”
- “Clearly, the defenda[illegible] ilty”

WRONG !

The Prosecutor's Fallacy - Example

Let’s recall the statement of the expert witness: “An innocent person has a 1 in 100,000 chance of matching the damning evidence (DE)”

This implies: $P(\mathrm{DE|I}) = 1/100{,}000 = 0.001\%$

The Prosecutor’s Fallacy is to conclude that:

$P(I|DE) = P(\mathrm{DE|I}) = 0.001\%$

However, it is most likely that $P(I|DE) \neq P(\mathrm{DE|I})$

The Prosecutor's Fallacy - Example

However, we don’t know $P(I|DE)$. But, fortunately, we know Bayes Rule!

$$P(I|DE) \ = P(DE|I) \times \ P(I)/P(DE)$$

$P(I|DE) = P(\mathrm{DE}|\mathrm{I})$ if $P(I) = P(DE)$

The Prosecutor's Fallacy - Example

Assumptions:

- The guilty person is among the 500,000 adults living around the crime scene.
- The guilty person also matches the DE

$P(I)$ = 499,999/500,000 = 0.999998 $P(I')$ = 1/500,000 = 0.000002 $P(\mathrm{DE}|\mathrm{I}')$ = 1

$$P(DE) = P(DE|I) \times P(I) + P(DE|I') \times P(I')$$
$$= 0.00001 \times 0.999998 + 1 \times 0.000002$$
$$= 0.000012$$

There are 5 innocent and 1 guilty person who match the DE

$$P(I|DE) = P(DE|I) \times P(I)/P(DE)$$
$$= 0.00001 \times 0.999998/0.000012$$
$$= 0.83333$$

5 out of 6 DE matches are innocent

$$P(I'|DE) = 1 - P(I|DE) = 0.16667$$

BUSINESS SCHOOL | RWTH AACHEN UNIVERSITY

Discrete Random Variables: Definitions

Random variable**:** A random variable assigns a single numerical value to each basic outcome in the sample space.

Discrete random variable: A discrete random variable X can take on a list of possible values ("finitely many" or "countable infinitely many")

Probability Distribution: A collection of probabilities for random variables

Cumulative Probability: Probability of a random variable X taking values smaller than x. E.g.:

$$F(x) = P(X \leq x)$$

Discrete Random Variables: Sum of Two Dice - Probabilities

$P(X = 2) = 1/36$

$P(X = 3) = 2/36$

$P(X = 7) = 6/36$

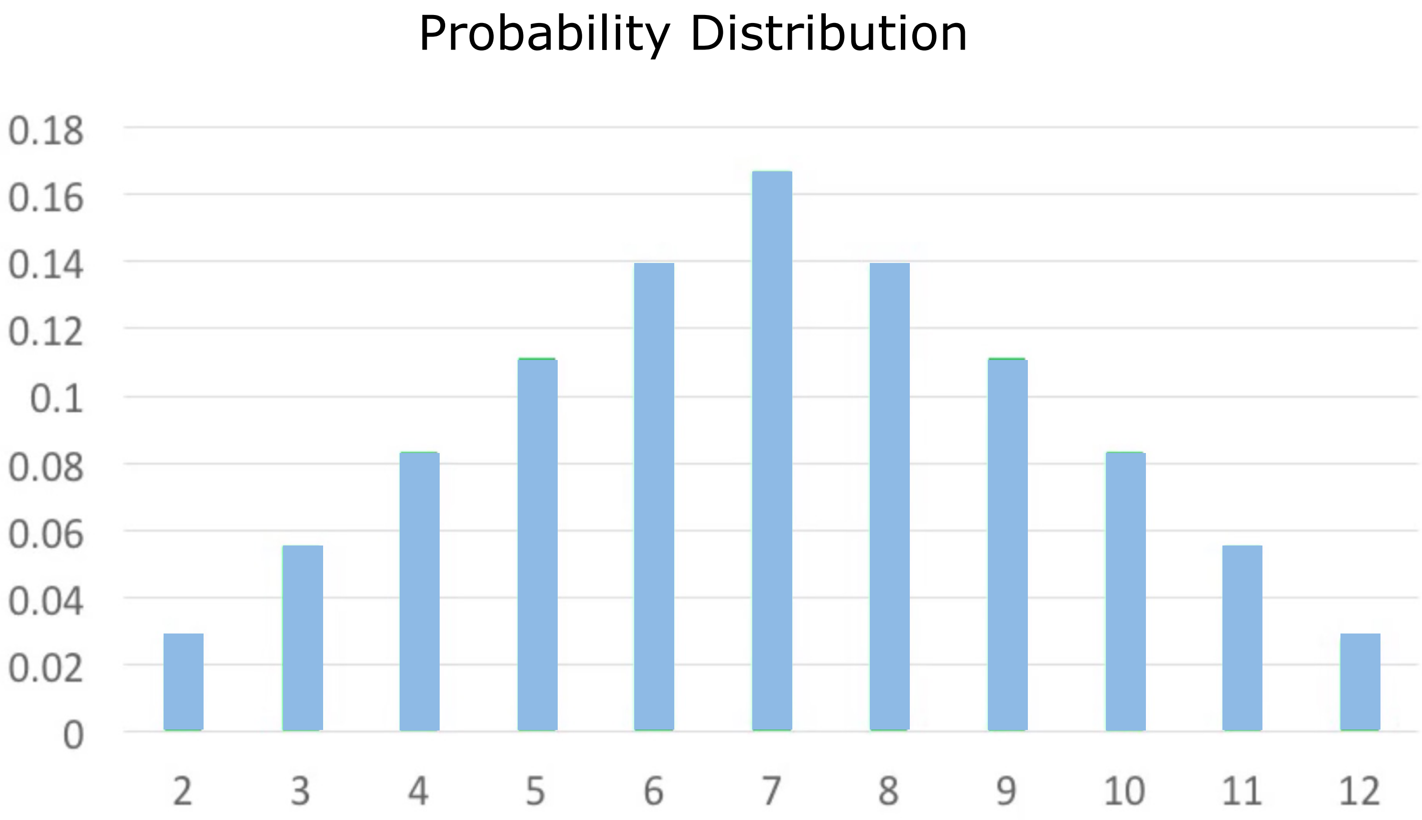
$P(X = 12) = 1/36$

X	P(X=x)
2	0.027778
3	0.055556
4	0.083333
5	0.111111
6	0.138889
7	0.166667
8	0.138889
9	0.111111
10	0.083333
11	0.055556
12	0.027778

$P(X = x) = 0$ for all values x not in {2, 3, 4, …, 12}

Discrete Random Variables: Sum of Two Dice - Distribution

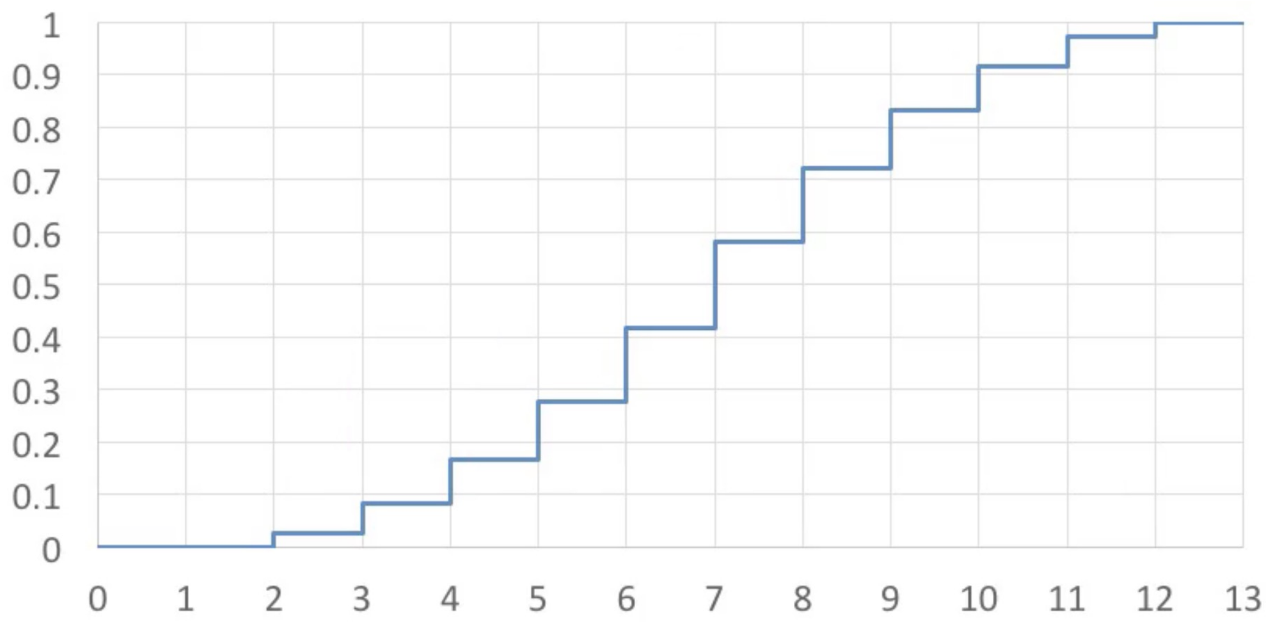
X	P(X=x)
2	0.027778
3	0.055556
4	0.083333
5	0.111111
6	0.138889
7	0.166667
8	0.138889
9	0.111111
10	0.083333
11	0.055556
12	0.027778



Discrete Random Variables: Sum of Two Dice – Cumulative Distribution

X	P(X=x)	F(x)
2	0.027778	0.027778
3	0.055556	0.083333
4	0.083333	0.166667
5	0.111111	0.277778
6	0.138889	0.416667
7	0.166667	0.583333
8	0.138889	0.722222
9	0.111111	0.833333
10	0.083333	0.916667
11	0.055556	0.972222
12	0.027778	1

Cumulative Distribution



Expected Value

As soon as there are many possible outcomes, we want to summarize the information on random variables into summary measures.

The basic summary measure is the mean, or the expected value.

Expected value of a random variable: Mean of a discrete random variable X is the probability-weighted sum of all possible values.

$$\mu = E(X) = x_1 p(x_1) + x_2 p(x_2) + \ldots + x_k p(x_k)$$

Expected Value: 10- Sided Die

D	P(D=d)	d*P(d)
1	0,1	0,1
2	0,1	0,2
3	0,1	0,3
4	0,1	0,4
5	0,1	0,5
6	0,1	0,6
7	0,1	0,7
8	0,1	0,8
9	0,1	0,9
10	0,1	1
Sum	1	5,5
		E(D)

Expected Value: Betting on "Red" in Roulette

Suppose you bet on red:

- Outcome = red: net win of 1
- Outcome = black/green: net loss of 1

The random variable is B, with possible outcomes 1 and -1

$$P(B = 1) \quad = 18/37 = 0.486$$

$$P(B = -1) = 19/37 = 0.514$$

$$E(B) = 1 \times 18/37 + (-1) \times 19/37 = -1/37 = -0.027$$

The Bank always wins… in the long run!

Expected Value: Should You Buy the Extra Insurance?

Suppose your current bike theft insurance has a deductible amount of 200€. The company offers you to eliminate that deductible amount entirely with an extra insurance package that costs 40€

Will you buy the extra insurance package?

Expected Value: Should You Buy the Extra Insurance?

Suppose your current bike theft insurance has a deductible amount of 200€. The company offers you to eliminate that deductible amount entirely with an extra insurance package that costs 40€

Will you buy the extra insurance package?

Without the extra insurance:

B_1 has two values, 0 (no theft) and 200 (theft)

The expected payment is:

$$E(B_1) = 200 \times p + 0 \times (1-p) = 200 \times p$$

Wit the extra insurance:

B_2 has two values, 0 (no theft) and 0 (theft)

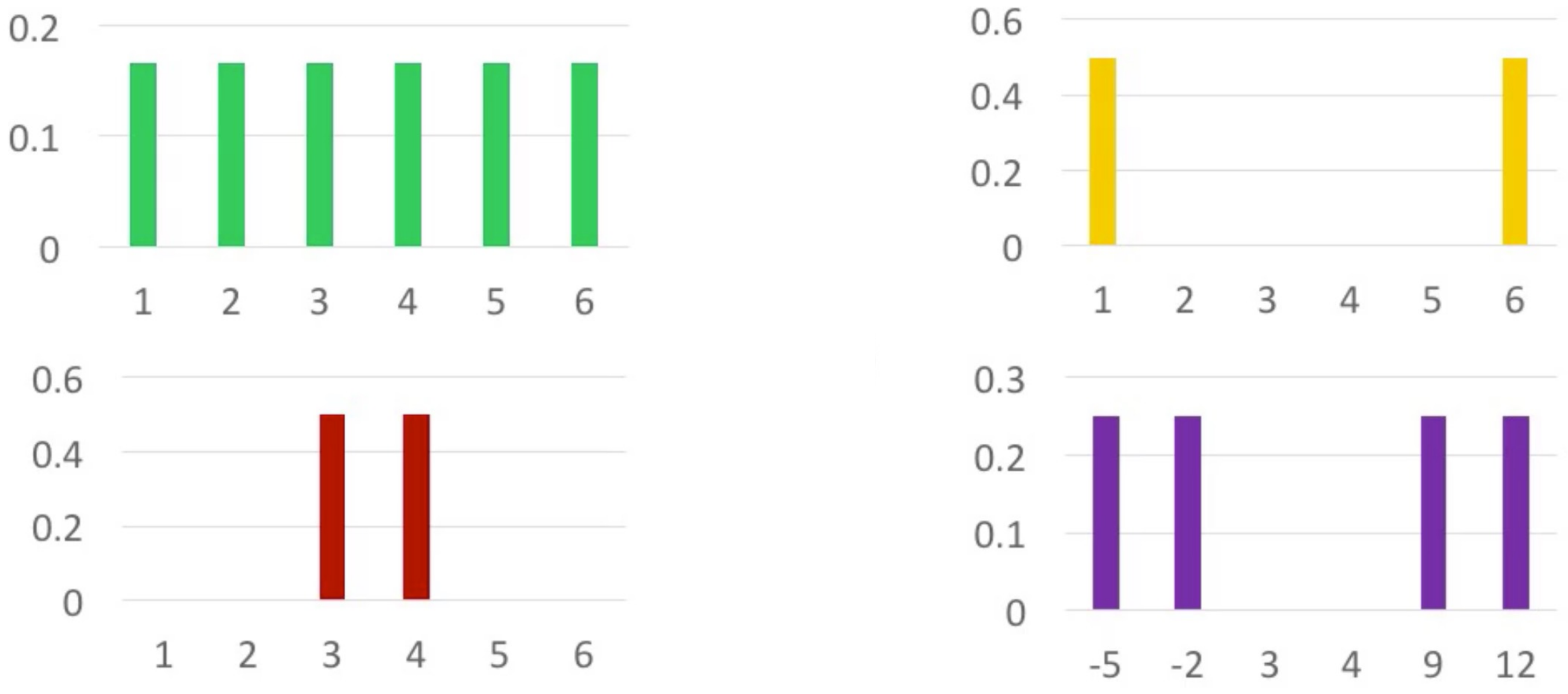
The expected payment is:

$$E(B_2) + 40 = 0 \times p + 0 \times (1-p) + 40 = 40$$

We have no idea what the probability of theft is. Since, for $p \leq 0.2$, the insurance company is making money, the insurance company must believe that $p \leq 0.2$.

Measures of Variation

The expected value entails no information about variation. Consider the following four distributions, each having an expected value of 3.5

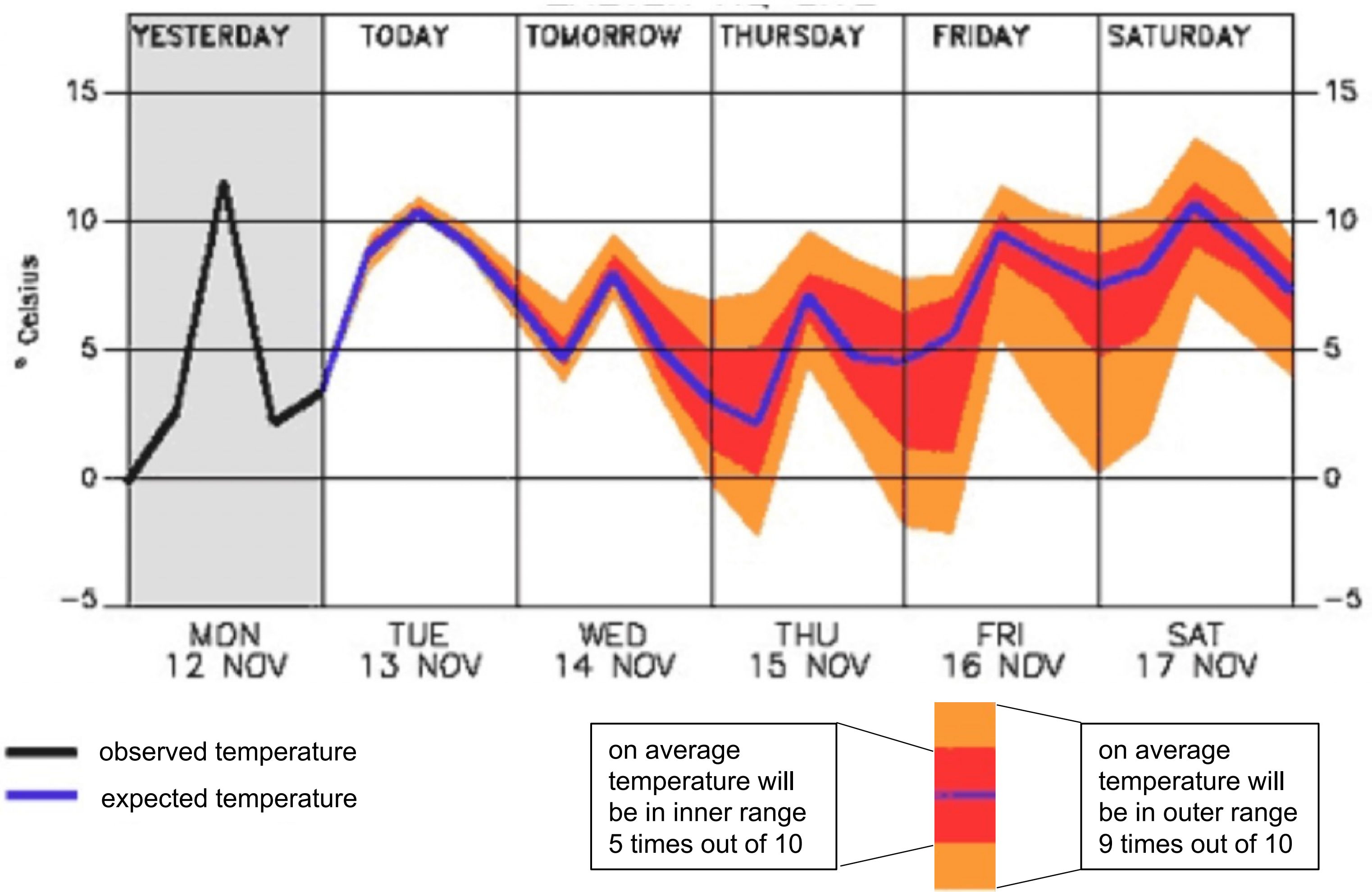


Measures of Variation

People rely too often simply on expected values

Measures of Variation

However, behind expected values is always some information about variation.



BUSINESS SCHOOL | RWTHAACHEN UNIVERSITY

Measures of Variation: Variance

Variance: The variance is a probability-weighted sum of the squared deviations from the mean.

Suppose X has k possible values, $x_1, x_1, \ldots, x_k,$ then the variance of X is

$$
\begin{aligned}
\sigma^2 &= Var(X) \\
&= E((X-\mu)^2) \\
&= (x_1 - \mu)^2 p(x_1) + (x_2 - \mu)^2 p(x_2) + \ldots + (x_k - \mu)^2 p(x_k)
\end{aligned}
$$

Measures of Variation: Standard Deviation

The units of variance is also the unit of the variable squared. We want to measure variation in the same unit as we measure the variable.

Standard Deviation: is a measure (based on the variance) for the average deviation of the values of a random variable X from its mean μ.

$$\sigma = SD(X) = \sqrt{Var(X)}$$

Measures of Variation: 10-Sided Die

D	$P(D = d)$	$d \times P(d)$	$d - E(D)$	$(d - E(D))^2 \times P(d)$	
1	0.1	0.1	-4.5	2.025	
2	0.1	0.2	-3.5	1.225	
3	0.1	0.3	-2.5	0.625	
4	0.1	0.4	-1.5	0.225	
5	0.1	0.5	-0.5	0.025	
6	0.1	0.6	0.5	0.025	
7	0.1	0.7	1.5	0.225	
8	0.1	0.8	2.5	0.625	
9	0.1	0.9	3.5	1.225	
10	0.1	1	4.5	2.025	
Sum	1	5.5		8.25	2.8722813
		E(D)		**Var(D)**	**SD(D)**

BUSINESS SCHOOL | RWTH AACHEN UNIVERSITY

Bernoulli Random Variable

Bernoulli random variable B allows for only two outcomes, b=1 (success) and b=0 (failure).

$$P(B = 1) = p$$

$$P(B = 0) = 1 - p$$

Bernoulli Random Variable

Bernoulli random variable B allows for only two outcomes, b=1 (success) and b=0 (failure).

$$P(B = 1) = p$$

$$P(B = 0) = 1 - p$$

Mean of Bernoulli random variable B:

$$\mathrm{E}(B) = \mu = 1 \times p + 0 \times (1 - p) = p$$

Variance of Bernoulli random variable B:

$$\begin{aligned}\mathrm{Var}(B) &= \sigma^2 \\ &= (1 - p)^2 \times p + (0 - p)^2 \times (1 - p) \\ &= p(1 - p)\end{aligned}$$

Binomial Trails

A sequence of Bernoulli random variables is called Bernoulli trials:

Let X be a binomial random variable:

$$X = B_1 + B_2 + \ldots + B_n$$

X has the binomial distribution B(n,p)

Mean: $\mathrm{E}(X) = E(B_1 + B_2 + \ldots + B_n)$

$$= p + p + \cdots + p$$

$$= np$$

Variance: $\mathrm{Var}(X) = \mathrm{Var}(B_1 + B_2 + \ldots + B_n)$

$$= p(1-p) + p(1-p) + \cdots + p(1-p)$$

$$= np(1-p)$$

Binomial Distribution: Example – Repeated Flips of a Coin

Y = number of heads (H)

Y = 0: 5 Tails (T), i.e.: TTTTT

$$P(Y = 0) = 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 = 1/32 = 0.03125$$

Y = 1: HTTTT, THTTT, TTHTT, TTTHT, TTTTH

$$P(Y = 1) = 5/32 = 0.15625$$

Y	P(Y=y)	F(y)
0	0.03125	0.03125
1	0.15625	0.1875
2	0.3125	0.5
3	0.3125	0.8125
4	0.15625	0.96875
5	0.03125	1

Binomial Distribution: Example – Repeated Flips of a Coin

Five coin flips (in Excel)

P(Y=0) = BINOM.DIST(0,5,0.5,FALSE) = 0.03125
P(Y=1) = BINOM.DIST(1,5,0.5,FALSE) = 0.15625
P(Y=2) = BINOM.DIST(2,5,0.5,FALSE) = 0.3125
P(Y=3) = BINOM.DIST(3,5,0.5,FALSE) = 0.3125
P(Y=4) = BINOM.DIST(4,5,0.5,FALSE) = 0.15625
P(Y=5) = BINOM.DIST(5,5,0.5,FALSE) = 0.03125

Number of successes

Number of trials

Probability of success

Show cumulative distribution?

Y	P(Y=y)	F(y)
0	0.03125	0.03125
1	0.15625	0.1875
2	0.3125	0.5
3	0.3125	0.8125
4	0.15625	0.96875
5	0.03125	1

Binomial Distribution: Application – Schmitz Live Beer Advertisement

In 1981, Schlitz Brewing Company conducted a live beer testing during the halftime of the Super Bowl. 100 beer drinkers who favored Michelob (a direct competitor) would try blindly both beers and had (live on TV!) to decide which one they would favor. This has been considered a risky and shocking marketing move.

Was is really that risky? Here are a few quite realistic assumptions:

- The Schlitz beer and the Michelob beer and in fact many other beers tasted quite the same.
- Essentially this is like coin flipping.

Binomial Distribution: Application – Schmitz Live Beer Advertisement

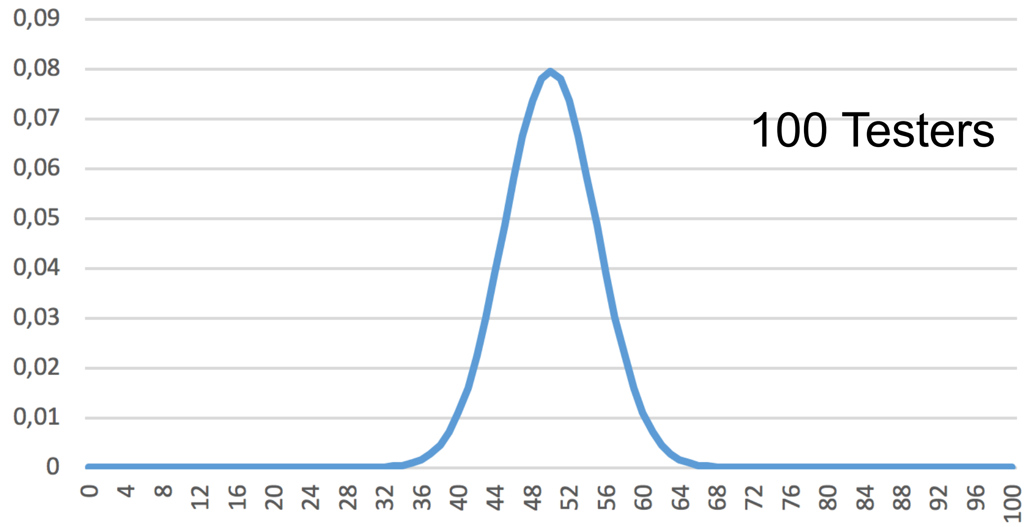
This reduces to a binomial experiment (Bernoulli trail):

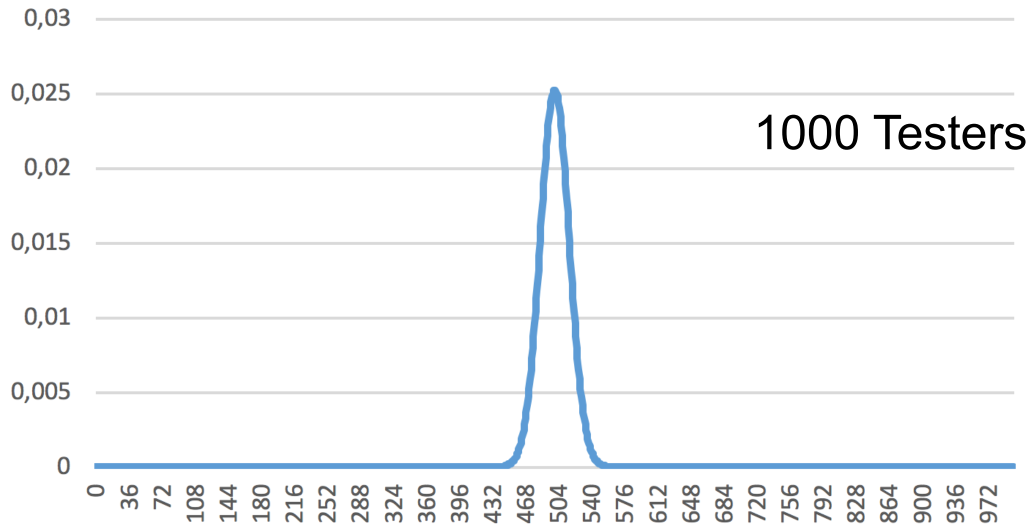
- Fixed number of trials (100 testers)
- Two possible outcomes (Michelob or Schlitz)
- Success: Tester picking Schlitz
- Probability of success the same for each trail
- Trails are independent (testers do not influence each other)

The probability that all 100 testers would prefer the competitor‘s beer was $\frac{1}{7.88 \cdot 10^{32}}$

The probability that at least 45 testers would prefer Schlitz was 84%

Binomial Distribution: Application – Schmitz Live Beer Advertisement





BUSINESS SCHOOL | RWTH AACHEN UNIVERSITY

Binomial Distribution: Application

Binomial Distribution: Application

https://www.youtube.com/watch?v=P9a3K2vkvrU

Binomial Distribution: Application – Airlines Overbooking

Airlines tend to overbook their flights in order to increase revenues

How do airlines make the tradeoff between higher revenues and loss of customer satisfaction and cost of compensation?

- All people make show-up decisions independently with constant probability p
- We can apply the binomial distribution

Move to Excel

RWTH BUSINESS SCHOOL

Mathematics & Statistics
M.Sc. Data Analytics and Decision Science
Prof. Dr. Thomas S. Lontzek

AACSB ACCREDITED

BUSINESS SCHOOL | RWTH AACHEN UNIVERSITY