Statistical Methods and Data Analysis I
Lecture 4: Probability Basics

Oleg Goldshmidt
oleg.goldshmidt@post.idc.ac.il

Arison School of Business
Interdisciplinary Center (IDC)
Herzliya, Israel

March 12, 2019

Confusion With Confusion Tables

Testing H_0	(match) Reject H_0	(no match) Do not reject H_0
H_0 false (guilty)	True positive (correct decision)	False negative (Type II error)
H_0 true (innocent)	False positive (Type I error)	True negative (correct decision)

Testing H_0	(no match) Do not reject H_0	(match) Reject H_0
H_0 true (innocent)	True negative (correct decision)	False positive (Type I error)
H_0 false (guilty)	False negative (Type II error)	True positive (correct decision)

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.

How rational is this policy?

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.

How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test?

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test? Very — better than most on the market. It determines drug users with 99% accuracy.

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test? Very — better than most on the market. It determines drug users with 99% accuracy. False positive rate is 1%

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test? Very — better than most on the market. It determines drug users with 99% accuracy. False positive rate is 1%
2. And what is the *a priori* probability that a randomly selected person will be a drug user?

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test? Very — better than most on the market. It determines drug users with 99% accuracy. False positive rate is 1%
2. And what is the *a priori* probability that a randomly selected person will be a drug user? Assume 1 out of 1000

Example: Drug Screening in Workplace

Scenario:

A large firm's HR department decided to subject every job applicant to urinanalysis for a range of illegal drugs. Anyone who tests positive is rejected as a matter of policy.
How rational is this policy?

Aside: is this scenario realistic? It was, at one time: "We need to have drug testing in most areas of work." [Edwin Meese, USAG, mid-80ies]. Many Fortune 500 companies used on-the-job tests and pre-employment screening.

1. How accurate is the test? Very — better than most on the market. It determines drug users with 99% accuracy. False positive rate is 1%
2. And what is the *a priori* probability that a randomly selected person will be a drug user? Assume 1 out of 1000
3. The company has 1000 applicants for various positions.

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user?

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	2	998
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	2	998
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.
- And what if 1% of population are drug users?

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.
- And what if 1% of population are drug users? Well: fifty-fifty then.

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.
- And what if 1% of population are drug users? Well: fifty-fifty then.
- Are drug tests useless???

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.
- And what if 1% of population are drug users? Well: fifty-fifty then.
- Are drug tests useless???
- No, not if tests are administered to those who manifest drug-induced behavior and when all positive tests are repeated (as sports doping agencies do). Indiscriminate tests are impossible to interpret meaningfully, though.

Example: Drug Screening in Workplace — Analysis

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

- A positive test corresponds to about 9.1% (1/11) chance that the applicant is a drug user.
- What if only 1 in 500 is a drug user? A positive test corresponds to 16.7% (2/12) probability of correct determination.
- And what if 1% of population are drug users? Well: fifty-fifty then.
- Are drug tests useless???
- No, not if tests are administered to those who manifest drug-induced behavior and when all positive tests are repeated (as sports doping agencies do). Indiscriminate tests are impossible to interpret meaningfully, though.

Aside: when such tests were widespread legislation was proposed to forbid dismissals on the basis of a single test. However, such tests are expensive, and there was never any protection for job applicants.

Example: Police Face Recognition Trials

Scenario

UK Police ran a number of trial deployments — Notting Hill (2017), Soho, Piccadilly Circus, Leicester Sq. (2018), Romford (2019) — of face recognition software, scanning faces of passersby and matching them against a watch list.

Example: Police Face Recognition Trials

Scenario

UK Police ran a number of trial deployments — Notting Hill (2017), Soho, Piccadilly Circus, Leicester Sq. (2018), Romford (2019) — of face recognition software, scanning faces of passersby and matching them against a watch list.

- Widespread criticism on privacy, efficacy grounds.
- 0 arrests, 2 correct IDs, 98% false positives.
- This does **not** mean that 98% of passersby are wrongly flagged up. Instead, it means that out of those who **have been** flagged up 98% were **not** on the watch list.

Example: Police Face Recognition Trials

Scenario

UK Police ran a number of trial deployments — Notting Hill (2017), Soho, Piccadilly Circus, Leicester Sq. (2018), Romford (2019) — of face recognition software, scanning faces of passersby and matching them against a watch list.

- *Assume 99% efficacy of the technology* (p-value = 0.01).
- Assume 9800 faces scanned in a trial deployment.

- Widespread criticism on privacy, efficacy grounds.
- 0 arrests, 2 correct IDs, 98% false positives.
- This does **not** mean that 98% of passersby are wrongly flagged up. Instead, it means that out of those who **have been** flagged up 98% were **not** on the watch list.

Example: Police Face Recognition Trials

Scenario

UK Police ran a number of trial deployments — Notting Hill (2017), Soho, Piccadilly Circus, Leicester Sq. (2018), Romford (2019) — of face recognition software, scanning faces of passersby and matching them against a watch list.

- *Assume 99% efficacy of the technology* (p-value = 0.01).
- Assume 9800 faces scanned in a trial deployment.

- Widespread criticism on privacy, efficacy grounds.
- 0 arrests, 2 correct IDs, 98% false positives.
- This does **not** mean that 98% of passersby are wrongly flagged up. Instead, it means that out of those who **have been** flagged up 98% were **not** on the watch list.

Hypothesis	Match	No match
H_A (suspect)	2	9798
H_0 (innocent)	98	9702

Random Process and Random Variables

Random process

A situation where we know what the possible outcomes are, but not which particular outcome will occur in each specific observation.

Random variable

Suppose we measure the value of a variable in a series of observations or experiments. If the measure value is determined by some *random process* then we say that the variable we measure is a random variable.

Another term: stochastic.

Examples: coin toss, dice roll, playlist shuffle, stock market prices, currency exchange rates, election poll responses, sales numbers, factory production defects, sick days taken by employees, student grades, time to get to work every morning, number of visitors to a web page, number of Instagram selfies posted per unit time...

Does not have to be truly random...

Probability

- The notion of *probability* is essential for describing random variables.
- **Definition:** Probability is the measure of the likelihood that an event will occur.
  - NB: the definition is informal. Formally, "likelihood" and "probability" are distinct notions, but we will ignore that for now.
  - "Event" may mean an outcome of an observation or an experiment.
  - Applies to numerical and categorical variables.
- Various definitions and interpretations (almost) completely agree on the mathematical rules probability must follow.
- **Notation:** probability of event (outcome) A is $P(A)$.
- **By definition:** $0 \leq P(A) \leq 1$.
  - Often expressed as a percentage.

Objectivist Interpretation of Probability

Probability reflects some objective state of things.
There is no preference for a given view or hypothesis.

Frequentist interpretation

$P(A)$ is the relative proportion of times outcome A will occur if we repeat an observation or experiment, “in the long run”.
A bit more formally, $P(A)$ is the proportion of times outcome A would occur if we observed the underlying random process and measured the values of the relevant random variables an *infinite* number of times.

$$P(A) = \lim_{N\to\infty} \frac{N_A}{N}$$

where N is the total number of measurements, N_A is the number of times A occurred.

Law of Large Numbers

Law of Large Numbers

As more observations are collected, the proportion of occurrences of a particular outcome A converges to the probability of that outcome $P(A)$.

Law of Large Numbers

Law of Large Numbers

As more observations are collected, the proportion of occurrences of a particular outcome A converges to the probability of that outcome $P(A)$.

- When tossing a fair coin heads came up 10 times in a row. What is the probability of getting heads on the 11th toss?

$$HHHHHHHHHH?$$

Law of Large Numbers

Law of Large Numbers

As more observations are collected, the proportion of occurrences of a particular outcome A converges to the probability of that outcome $P(A)$.

- When tossing a fair coin heads came up 10 times in a row. What is the probability of getting heads on the 11th toss?

$$HHHHHHHHHH?$$

- *No, the coin is not "due" for a tails!* The probability is still 0.5.

Law of Large Numbers

As more observations are collected, the proportion of occurrences of a particular outcome A converges to the probability of that outcome $P(A)$.

- When tossing a fair coin heads came up 10 times in a row. What is the probability of getting heads on the 11th toss?

$$HHHHHHHHHH?$$

- *No, the coin is not "due" for a tails!* The probability is still 0.5.
- *Gambler's fallacy* is a misunderstanding of the Law of Large Numbers: *in general* random processes do not "compensate" for whatever happened in the past.
  - Discussion follows...

Disjoint and Non-Disjoint Outcomes

Disjoint outcomes

Mutually exclusive events that cannot happen at the same time.

- A coin toss can be either heads or tails.
- A student cannot pass and fail a class at the same time.
- For disjoint events A and B

$$P(A \text{ and } B) = 0, \quad P(A \text{ or } B) = P(A) + P(B)$$

Non-disjoint outcomes

Non-exclusive events that can happen at the same time.

- A card may be a jack and a red.
- A student can get a perfect grade in both Statistical Methods and Linear Algebra.
- For non-disjoint events A and B

$$P(A \text{ and } B) > 0, \quad P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Disjoint and Non-Disjoint Examples

Disjoint and Non-Disjoint Examples

- *What is the probability to roll either 5 or 6 on a fair dice?*

Disjoint and Non-Disjoint Examples

- *What is the probability to roll either 5 or 6 on a fair dice?*

$$P(5 \text{ } or \text{ } 6) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Disjoint and Non-Disjoint Examples

- *What is the probability to roll either 5 or 6 on a fair dice?*

$$P(5\ or\ 6) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

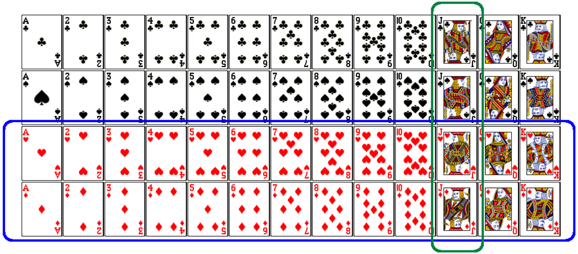
- *What is the probability to draw a red card or a jack from a well-shuffled deck?*

Disjoint and Non-Disjoint Examples

- *What is the probability to roll either 5 or 6 on a fair dice?*

$$P(5 \text{ or } 6) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

- *What is the probability to draw a red card or a jack from a well-shuffled deck?*

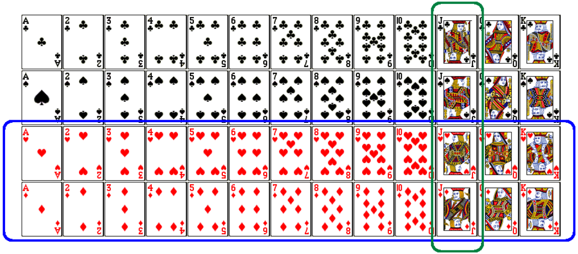


Disjoint and Non-Disjoint Examples

- *What is the probability to roll either 5 or 6 on a fair dice?*

$$P(5 \text{ or } 6) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

- *What is the probability to draw a red card or a jack from a well-shuffled deck?*



$$P(R \text{ or } J) = P(R) + P(J) - P(R \text{ and } J) = \frac{26}{52} + \frac{4}{52} - \frac{2}{52} = \frac{7}{13}$$

Independent and Dependent Processes and Events

Independent processes:

Knowing the outcome of one does not provide any information about the outcome of the other.

Independent variables:

Knowing the value of one does not provide any information about the value of the other.

Independent and Dependent Processes and Events

Independent processes:

Knowing the outcome of one does not provide any information about the outcome of the other.

Independent variables:

Knowing the value of one does not provide any information about the value of the other.

- If events A and B are independent then

$$P(A \text{ } and \text{ } B) = P(A) \times P(B)$$

Independent and Dependent Processes and Events

Independent processes:

Knowing the outcome of one does not provide any information about the outcome of the other.

Independent variables:

Knowing the value of one does not provide any information about the value of the other.

- If events A and B are independent then

$$P(A \text{ and } B) = P(A) \times P(B)$$

- Coin toss:
  - probability of heads: $P(H) = 1/2$
  - probability of getting heads twice in a row: $P(2 \times H) = (1/2) \times (1/2) = 1/4$
  - probability of getting heads 10 times in a row: $P(10 \times H) = (1/2)^{10} = 1/1024$

Independent and Dependent Processes and Events

Independent processes:

Knowing the outcome of one does not provide any information about the outcome of the other.

Independent variables:

Knowing the value of one does not provide any information about the value of the other.

- If events A and B are independent then

$$P(A \textit{ and } B) = P(A) \times P(B)$$

- Coin toss:
  - probability of heads: $P(H) = 1/2$
  - probability of getting heads twice in a row: $P(2 \times H) = (1/2) \times (1/2) = 1/4$
  - probability of getting heads 10 times in a row: $P(10 \times H) = (1/2)^{10} = 1/1024$
- If A and B are *not* independent then we need to discuss *conditional probabilities*:
  - probability of A given B: $P(A|B)$, probability of B given A: $P(B|A)$

Conditional Probability

$$P(A \; and \; B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

Conditional Probability

$$P(A \text{ and } B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

- *What is the probability that* $1, 2, 3, 4, 5, 6$ *will be drawn in the "6 out of 37" lottery?*

Conditional Probability

$$P(A \text{ and } B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

- *What is the probability that $1, 2, 3, 4, 5, 6$ will be drawn in the "6 out of 37" lottery?*

$$P(6/37) = \frac{6}{37} \times \frac{5}{36} \times \frac{4}{35} \times \frac{3}{34} \times \frac{2}{33} \times \frac{1}{32} \approx 0.00000043$$

Conditional Probability

$$P(A \text{ and } B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

- *What is the probability that* $1, 2, 3, 4, 5, 6$ *will be drawn in the "6 out of 37" lottery?*

$$P(6/37) = \frac{6}{37} \times \frac{5}{36} \times \frac{4}{35} \times \frac{3}{34} \times \frac{2}{33} \times \frac{1}{32} \approx 0.00000043$$

(this is a chained application of the conditional probability formula)

Conditional Probability

$$P(A \textit{ and } B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

- *What is the probability that* $1, 2, 3, 4, 5, 6$ *will be drawn in the "6 out of 37" lottery?*

$$P(6/37) = \frac{6}{37} \times \frac{5}{36} \times \frac{4}{35} \times \frac{3}{34} \times \frac{2}{33} \times \frac{1}{32} \approx 0.00000043$$

(this is a chained application of the conditional probability formula)

- Computing conditional probability:

$$P(A|B) = \frac{P(A \textit{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \textit{ and } B)}{P(A)}$$

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A \; and \; B)}{P(B)}, \quad P(B|A) = \frac{P(A \; and \; B)}{P(A)}$$

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A\ and\ B)}{P(B)}, \quad P(B|A) = \frac{P(A\ and\ B)}{P(A)}$$

$$P(A\ and\ B) = P(A) \times P(B)$$

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

$$P(A \text{ and } B) = P(A) \times P(B)$$

$$P(A|B) = P(A), \quad P(B|A) = P(B)$$

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A\ and\ B)}{P(B)}, \quad P(B|A) = \frac{P(A\ and\ B)}{P(A)}$$

$$P(A\ and\ B) = P(A) \times P(B)$$

$$P(A|B) = P(A), \quad P(B|A) = P(B)$$

This explains the Gambler's fallacy:

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A \textit{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \textit{ and } B)}{P(A)}$$

$$P(A \textit{ and } B) = P(A) \times P(B)$$

$$P(A|B) = P(A), \quad P(B|A) = P(B)$$

This explains the Gambler's fallacy:

$$P(H|10 \times H) = P(H) = \frac{1}{2}, \quad P(T|10 \times H) = P(T) = \frac{1}{2}$$

Conditional Probability For Independent Variables

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

$$P(A \text{ and } B) = P(A) \times P(B)$$

$$P(A|B) = P(A), \quad P(B|A) = P(B)$$

This explains the Gambler's fallacy:

$$P(H|10 \times H) = P(H) = \frac{1}{2}, \quad P(T|10 \times H) = P(T) = \frac{1}{2}$$

Stochastic processes (coin toss, roulette, markets) can remain irrational longer than you can remain solvent...

Bayes Theorem

Our formulas for conditional probabilities:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

Our formulas for conditional probabilities:

$$P(A|B) = \frac{P(A \text{ } and \text{ } B)}{P(B)}, \quad P(B|A) = \frac{P(A \text{ } and \text{ } B)}{P(A)}$$

Bayes Theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}, \quad P(A|B) = \frac{P(B|A)}{P(B)} \times P(A)$$

Bayes Theorem

Our formulas for conditional probabilities:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}, \quad P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

Bayes Theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}, \quad P(A|B) = \frac{P(B|A)}{P(B)} \times P(A)$$

- Terminology:
  - $P(A|B)$, $P(B|A)$ — *conditional probabilities*
  - $P(A)$, $P(B)$ — *marginal probabilities*

IDC HERZLIYA

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)}
\end{aligned}
$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)} \\
&= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.999}
\end{aligned}
$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	1	999
Drug-free	10	990

$$P(User|Positive) = \frac{P(Positive|User) \times P(User)}{P(Positive)}$$

$$= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)}$$

$$= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.999}$$

$$\approx 0.09$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	2	998
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)} \\
&= \frac{0.99 \times 0.002}{0.99 \times 0.002 + 0.01 \times 0.998}
\end{aligned}
$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	2	998
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)} \\
&= \frac{0.99 \times 0.002}{0.99 \times 0.002 + 0.01 \times 0.998} \\
&\approx 0.166
\end{aligned}
$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)} \\
&= \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.01 \times 0.99}
\end{aligned}
$$

Applying Bayes Theorem: Drug Screening

Hypothesis	Positive test	Negative test
Drug user	10	990
Drug-free	10	990

$$
\begin{aligned}
P(User|Positive) &= \frac{P(Positive|User) \times P(User)}{P(Positive)} \\
&= \frac{P(Positive|User) \times P(User)}{P(Positive|User) \times P(User) + P(Positive|Clean) \times P(Clean)} \\
&= \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.01 \times 0.99} \\
&\approx 0.5
\end{aligned}
$$

Applying Bayes Theorem: Police Face Recognition Trials

Hypothesis	Match	No match
H_A (suspect)	2	9798
H_0 (innocent)	98	9702

Notation: $S = suspect$, $I = innocent$, $M = match$

$$P(S|M) = \frac{P(M|S) \times P(S)}{P(M)}$$

$$= \frac{P(M|S) \times P(S)}{P(M|S) \times P(S) + P(M|I) \times P(I)}$$

Applying Bayes Theorem: Police Face Recognition Trials

Hypothesis	Match	No match
H_A (suspect)	2	9798
H_0 (innocent)	98	9702

Notation: $S = suspect, I = innocent, M = match$

$$P(S|M) = \frac{P(M|S) \times P(S)}{P(M)}$$

$$= \frac{P(M|S) \times P(S)}{P(M|S) \times P(S) + P(M|I) \times P(I)}$$

$$= \frac{0.99 \times (2/9800)}{0.99 \times (2/9800) + 0.01 \times (9798/9800)}$$

Applying Bayes Theorem: Police Face Recognition Trials

Hypothesis	Match	No match
H_A (suspect)	2	9798
H_0 (innocent)	98	9702

Notation: $S = suspect$, $I = innocent$, $M = match$

$$
\begin{aligned}
P(S|M) &= \frac{P(M|S) \times P(S)}{P(M)} \\
&= \frac{P(M|S) \times P(S)}{P(M|S) \times P(S) + P(M|I) \times P(I)} \\
&= \frac{0.99 \times (2/9800)}{0.99 \times (2/9800) + 0.01 \times (9798/9800)} \\
&\approx 0.0198
\end{aligned}
$$

Applying Bayes Theorem: Production Defects

- A factory with 3 production lines: 1) old, 2) newer, and 3) newest.
- The newer line produces 50% more and the newest — 2.5 times more than the old one.
- RMA rate is 5% for the old line, 3% for the newer line, 1% for the newest line.
- *What is the probability that an RMAed item was produced on the newest production line?*

Applying Bayes Theorem: Production Defects

- A factory with 3 production lines: 1) old, 2) newer, and 3) newest.
- The newer line produces 50% more and the newest — 2.5 times more than the old one.
- RMA rate is 5% for the old line, 3% for the newer line, 1% for the newest line.
- *What is the probability that an RMAed item was produced on the newest production line?*
- *Let X_i denote the event that an item was produced on line i ($i = 1, 2, 3$).*

$$P(X_1) = 0.2, \;\; P(X_2) = 0.3, \;\; P(X_3) = 0.5$$

Applying Bayes Theorem: Production Defects

- A factory with 3 production lines: 1) old, 2) newer, and 3) newest.
- The newer line produces 50% more and the newest — 2.5 times more than the old one.
- RMA rate is 5% for the old line, 3% for the newer line, 1% for the newest line.
- *What is the probability that an RMAed item was produced on the newest production line?*
- *Let X_i denote the event that an item was produced on line i ($i = 1, 2, 3$).*

$$P(X_1) = 0.2, \;\; P(X_2) = 0.3, \;\; P(X_3) = 0.5$$

- *Let D denote the event that a randomly chosen item is defective.*

$$P(D|X_1) = 0.05, \;\; P(D|X_2) = 0.03, \;\; P(D|X_3) = 0.01$$

Applying Bayes Theorem: Production Defects

- A factory with 3 production lines: 1) old, 2) newer, and 3) newest.
- The newer line produces 50% more and the newest — 2.5 times more than the old one.
- RMA rate is 5% for the old line, 3% for the newer line, 1% for the newest line.
- *What is the probability that an RMAed item was produced on the newest production line?*
- *Let X_i denote the event that an item was produced on line i ($i = 1, 2, 3$).*

$$P(X_1) = 0.2, \;\; P(X_2) = 0.3, \;\; P(X_3) = 0.5$$

- *Let D denote the event that a randomly chosen item is defective.*

$$P(D|X_1) = 0.05, \;\; P(D|X_2) = 0.03, \;\; P(D|X_3) = 0.01$$

$$P(D) = \sum_{i=1}^{3} P(D|X_i) \times P(X_i) = 0.05 \times 0.2 + 0.03 \times 0.3 + 0.01 \times 0.5 = 0.024$$

Applying Bayes Theorem: Production Defects

- A factory with 3 production lines: 1) old, 2) newer, and 3) newest.
- The newer line produces 50% more and the newest — 2.5 times more than the old one.
- RMA rate is 5% for the old line, 3% for the newer line, 1% for the newest line.
- *What is the probability that an RMAed item was produced on the newest production line?*
- *Let X_i denote the event that an item was produced on line i ($i = 1, 2, 3$).*

$$P(X_1) = 0.2, \;\; P(X_2) = 0.3, \;\; P(X_3) = 0.5$$

- *Let D denote the event that a randomly chosen item is defective.*

$$P(D|X_1) = 0.05, \;\; P(D|X_2) = 0.03, \;\; P(D|X_3) = 0.01$$

$$P(D) = \sum_{i=1}^{3} P(D|X_i) \times P(X_i) = 0.05 \times 0.2 + 0.03 \times 0.3 + 0.01 \times 0.5 = 0.024$$

$$P(X_i|D) = \frac{P(D|X_3) \times P(X_3)}{P(D)} = \frac{0.01 \times 0.5}{0.024} = \frac{5}{24}$$