ALQUSOUR ACADEMY
اكاديمية القصور

Q.A.J.U.S.T

BIOSTATISTICS

#لجنة_كلية_الصيدلة

لطلبة الصيدلة والعلوم الطبية

Subject:

First Exam – Part Three

ALQUSOUR ACADEMY

خاص
الفصل الدراسي الأول
2016/2015

Prod. Date: 3/11/2015 Pages: 11 Price: 40

ساعات الدوام الرسمي
السبت - الخميس: 11:00 ظهراً - 12:00 ليلاً
الجمعة: 2:00 ظهراً - 12:00 ليلاً

اربد / بجانب اربيلا مول / سلام سنتر / الطابق الرابع
للإستفسار والتسجيل: 0795 33 99 34 / 0785 70 60 08
Like facebook.com/alqusouracademy.page

ملاحظة هامة: تعتمد المادة على السلايدات.

Basic Probability Concepts and Applications

Probability terms:

Probability: the chance that an event will occur.

Event: is a possible outcome of an experiment. Example, a pregnancy outcome was a boy. Here the event is the boy.

Simple event: an event that can be described by a single characteristic (ex. red). Probability of having a red card = $\frac{26}{52}$ (because there is 26 red card and the cards deck is 52).

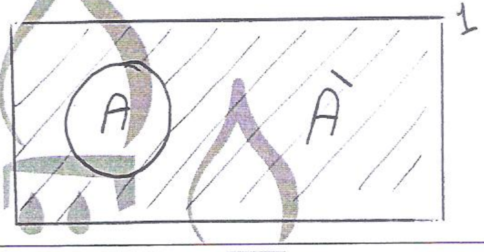
Sample space: is a collection of all possible events.
The sum of the probabilities of all the outcomes in the sample space is 1.

Mutually exclusive events: events that can not happen at same time. Thus, the probability of two mutually exclusive events occurring at the same time is zero. This means if A and B are mutually exclusive then their intersection = zero.

Independent events: two events A and B are independent if occurrence of A will not affect the occurrence of B, and vise versa.

Example:
A woman have girls in her 7 pregnancies. Then the probability of having a girl in her 8th pregnancy is 50% (it is not affected by the last pregnancy outcomes).

Complementary events: the complement of an event A consists of all events in the sample space not included in A, and denoted by $\bar{A}$. $So,\ P(A) = 1 - P(\bar{A})$



Probability Rules:
- The probability of any event (E) is a number between and including 0 and 1. This is denoted by:

$$0 \le P(E) \le 1$$

- If the event cannot occur (impossible), then its probability is 0.
- If the event is certain, then the probability is 1.

Some probability formulas:

(1) Marginal (**simple**) probability of A:

$$P(A) = \frac{total\ of(A)}{total}$$

It is the probability that an event happens at all, ignoring all other outcomes.

(2) **Joint** probability (multiplication rule):
It is the probability of two events happening simultaneously.

$$P(A \cap B) = P(A \text{ and } B) = \frac{A \text{ and } B \text{ intersection in the table}}{\text{total of all outcomes}}$$ (direct from the table).

(3) The **conditional** probability of A **given** B, is defined by:

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$$

Probability of B **given** A: $P\left(\frac{B}{A}\right) = \frac{P(A \cap B)}{P(A)}$

Note: You can find the intersection from this formula.

(4) The addition rule:
$P(A \cup B) = (A \text{ or } B) = P(A) + P(B) - P(A \cap B)$

(5) Suppose A and B are mutually exclusive events, then:
$P(A \cup B) = P(A) + P(B)$

(6) For **independent** events:

$P(A \cap B) = P(A)P(B)$

$P\left(\frac{A}{B}\right) = P(A)$

$P\left(\frac{B}{A}\right) = P(B)$

Example on independent:

A: get heads on the first toss, B: get heads on the second toss
Then P(A and B) = 0.5*0.5 = 0.25

Example:

A random sample of JUST students were selected to determine the smoking habit:

Specialty	Smoking (S)	Not smoking (N)	Total
Medical sciences (M)	20	40	60
Engineering (E)	24	72	96
Agriculture (A)	12	38	50
Humanities (H)	28	66	94
Total	80	220	300

If a student selected at random, find:

1) **The probability that the student is not smoker.**

$P(N) = \frac{220}{300} = 0.73$ (this is marginal probability)

2) **The probability that the student is specialized in Humanities.**

$P(H) = \frac{94}{300} = 0.31$ (this is marginal probability)

3) **The probability that he is agriculture and a smoking student .**

$P(A \cap S) = \frac{12}{300} = 0.04$

4) **The probability that he is agriculture student or not smoking.**

$P(A \cup N) = (A \text{ or } N) = P(A) + P(N) - P(A \cap N) =$

$$\frac{50}{300} + \frac{220}{300} - \frac{38}{300} = \frac{232}{300} = 0.77$$

5) **The probability that he is not an engineering student.**

$P(\bar{E}) = 1 - P(E) = 1 - \left(\frac{96}{300}\right) = 1 - 0.32 = 0.68$

6) **The probability that he is a medical science student if he is selected from smoking students.**

$$P\left(\frac{M}{S}\right) = \frac{P(M \cap S)}{P(S)} = \frac{\frac{20}{300}}{\frac{80}{300}} = \frac{20}{80} = 0.25$$

<u>Example:</u> Let $P(\bar{A}) = 0.4$, $P(B) = 0.7$, and $P(A \cap B) = 0.42$. Find the following:

(a) $P(A)$

$P(A) = 1 - P(\bar{A}) = 0.6$

(b) $P(A/B)$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = 0.6$$

(c) Decide if A and B are independent or not.

Since $P(A/B) = 0.6 = P(A)$, *they are independent*.

Applications of Probability-Screening Tests (Bays Theorem) ...

Screening: is the application of a test (or a symptom) to detect a potential disease in a person who has no known signs of that disease.
Usually done for general healthy public (community).

pharmacreative '14 ♡

Examples on screening:

- Mammogram.
- Pap smear.

In using a screening test we may get **A false positive** or **A false negative** result:

A false positive: results when a test indicates a positive status when the true status is negative. (positive but don't have the disease)
A false negative: results when a test indicates a negative status when the true status is positive. (negative but really have the disease)

	Disease		
Test Result	Present (D)	Absent ($\bar{D}$)	Total
Positive (T)	A	b	a + b
Negative ($\bar{T}$)	C	d	c + d
Total	a + c	b + d	n

b = # of false positives,
c = # of false negatives
a = # of true positives,
d = # of true negatives.

In order to test validity of a test before screening, we look for sensitivity and specificity.

- The **sensitivity** of a test (or symptom) is defined as:

The ability of the test to correctly detect those with the real disease.
So, it is the probability of a positive test result given the presence of the disease.

$$P(T/D) = \frac{a}{a + c}$$

- The **specificity** of a test (or symptom) is defined as:

The ability of the test to correctly identify those without the disease.
So, it is the probability of a negative result given the absence of the disease.

$$P(\bar{T}/\bar{D}) = \frac{d}{b + d}$$

Note: When sensitivity increases, specificity decreases (one increases, the other decreases)

- The **Prevalence** of disease (it is the probability of disease: it is always between 0 and 1) :

  Prevalence: measures the relative number of people in a population who have a particular disease at a given time. We write it P(D):

$$P(D) = \frac{\#\, existing\ cases}{total}$$

From the table it is: $\frac{a + c}{n}$

- **Incidence:** is the number of new cases of a disease that occurred during a specified period of time in a population at risk of developing the disease.

✓ The **predictive value positive** of a screening test (or symptom) is defined as:
The probability of having the disease given the test is positive.
(how likely someone with positive result actually have the disease).

✓ The **predictive value negative** of a screening test (or symptom) is defined by:
The probability of not having the disease when (given) the test is negative.
(how likely someone with negative test result actually don't have disease)

Formula from 2 × 2 tables:

- Positive predictive value = true positives / all positives. **(a/a + b)**
- Negative predictive value = true negatives / all negatives. **(d/c + d)**

Note: There is no perfect test with 100% sensitivity and 100% specificity.

We do screening for:

1) Common diseases (with high prevalence) .

   Because when disease prevalence increases, the predictive value positive increases and predictive value negative decreases.

   Important: sensitivity and specificity are not affected by prevalence.

2) Curable diseases.

Example:
A medical research team wishes to assess the usefulness of a certain symptom (call it S) in the diagnosis of particular disease. They use 2 groups. Results are in the following table:.

	Disease		
Test Result	Present (D)	Absent ($\bar{D}$)	Total
Positive (S)	744	21	765
Negative ($\bar{S}$)	31	1359	1390
Total	775	1380	2155

(a) What is the probability of a positive test result given that a patient has the disease? (sensitivity)

Solution:

$$P(S/D) = \frac{744}{775} = 0.96$$

(b) What is the probability of a negative test result given that a patient does not have the disease? (specificity)

Solution:

$$P(\bar{S}/\bar{D}) = \frac{1359}{1380} = 0.985$$

(c) What is the total number of false positives of this test? False negatives?

Solution: # of false positives = 21, and # of false negatives = 31.

(d) What is the disease prevalence?

Solution:

$$P(D) = \frac{\#\ existing\ cases}{total} = \frac{775}{2155} = 0.359$$

#لجنة كلية الصيدلة

(e) What is the predictive value positive of the symptom ?

Solution:

from the table:

predictive value positive $= \frac{true\ positives}{all\ positives} = \frac{744}{765} = 0.972$

Example :

Given the following table calculate sensitivity, specificity, predictive value positive , predictive value negative.

X-ray	Tuberculosis		
	No	**Yes**	**Total**
Negative	1739	8	1747
Positive	51	22	73
Total	1790	30	1820

Solution:

Sensitivity $= \frac{22}{30} = 0.733$

Specificity $= \frac{1739}{1790} = 0.972$

Positive Predictive Value $= \frac{22}{73} = 0.301$

Negative Predictive Value $= \frac{1739}{1747} = 0.995$
