Chapter 2 Probability

1. Definition of Probability
2. Probability of disjoint events
3. Probability of non-disjoint events
4. Probability of complement of an event
5. Probability of independent events
6. Probability of a conditional event
7. Probability of dependent events
8. Tree diagrams
9. Bayes Theorem
10. Mean of random variables
11. Variance of random Variables

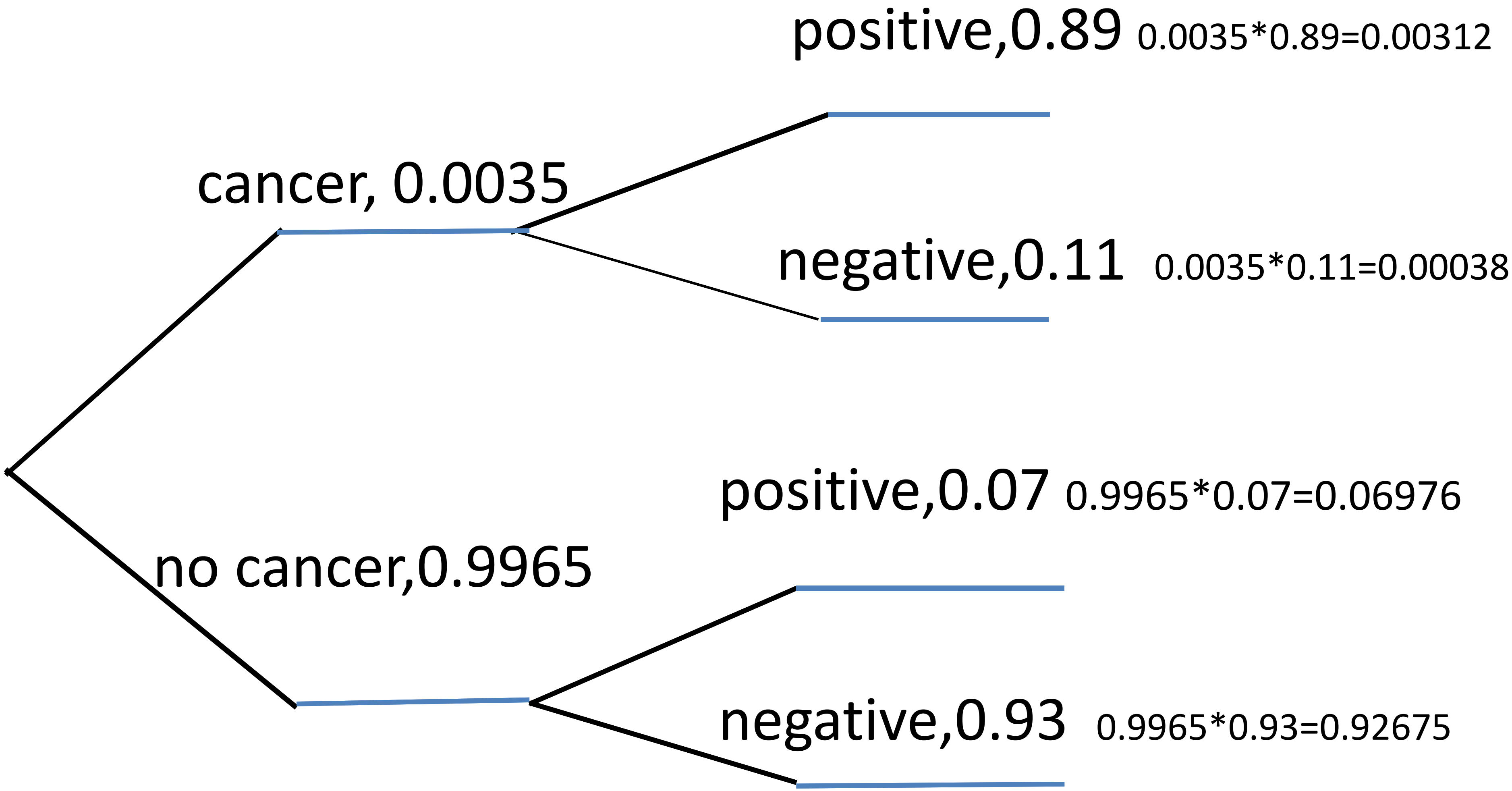
Example 1. In Canada, about 0.35% of women over 40 will be diagnosed with breast cancer in any given year. A common screening test for cancer is the mammogram, but this test is not perfect. In about 11% of patients with breast cancer, the test gives a false negative. Similarly, the test gives a false positive in 7% of patients who do not have breast cancer. If we tested a random woman over 40 for breast cancer using a mammogram and the test came back positive. What is the probability that the patient actually has breast cancer?

Answer: Our goal is to find P(has BC| test positive).

By the conditional probability

$$P(has\ \ BC \mid \quad test\ \ positive) = \frac{P(has\ \ BC\ \ and\ \ test\ \ positive)}{P(test\ \ positive)}$$

Tree diagram



Slide 3

1 Arthur Berg, 1/21/2015

From the tree diagram, we have

P(has BC and test positive) =0.00312

P(test positive)=P(has BC and test positive) +P(no BC and test positive)
=0.00312+0.06976
=0.07288

So P(has BC |test positive)=0.00312/0.07288=0.0428

Bayes Theorem

Conditional probability $P(A_1|B)$ can be calculated as

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \ldots + P(A_k)P(B \mid A_k)}$$

Let's verify the <u>**Example 1**</u> using Bayes Theorem.

$$P(\text{has BC} \mid \text{test postive}) = \frac{P(\text{has BC})P(\text{test positive} \mid \text{has BC})}{P(\text{has BC})P(\text{test positive} \mid \text{has BC}) + P(\text{no BC})P(\text{test positive} \mid \text{no BC})} = \frac{(0.0035)(0.89)}{(0.0035)(0.89)+(0.9965)(0.07)} = \frac{0.00312}{0.07288} = 0.0428$$

Mean and variance of random variables

Example 2: Two books are assigned for a statistics class: a textbook and its corresponding study guide. The university bookstore expected 20% of enrolled students do not buy either book, 55% buy the textbook, and 25% buy both books. The textbook cost $137 and the study guide $33. How much revenue should the book store expect from one student of this class?

Answer: Using basic arithmetic, we have

average revenue=137x55%+(137+33)x25%=117.85

What is the general formula in statistics for the mean of a random variable?

Mean of discrete random variables

	1	2	3	total
X_i	\$0	\$137	\$170	
P(**X**=X_i)	0.20	0.55	0.25	1.00

Expected value of a discrete random variable:
If x_1 , $x_2,$ x_3 , ……. . x_k **the all outcomes of a discrete variable X, then**

E(X)=x_1 **P(X=**x_1**)+**x_2 **P(X=**x_2**)+**x_3 **P(X=**x_3**)+ ……..**x_k**P(X=**x_k**)**

Answer: E(X)=0X0.20+137X0.55+170X0.25=117.85

<u>Example 3</u>: Two books are assigned for a statistics class: a textbook and its corresponding study guide. The university bookstore expected 20% of enrolled students do not buy either book, 55% buy the textbook, and 25% buy both books. The textbook cost $137 and the study guide $33. What is variance of the revenue from one student for the bookstore? What is the standard deviation of the revenue from one student for the bookstore?

Variance of discrete random variables

- General Variance Formula

$$\sigma^2 = (x_1 - \mu)^2 P(X = x_1) + (x_2 - \mu)^2 P(X = x_2) + (x_3 - \mu)^2 P(X = x_3) + \cdots\cdots (x_k - \mu)^2 P(X = x_k)$$

Here σ^2 is called variance and σ is called standard deviation.

	1	2	3	total
X_i	\$0	\$137	\$170	
P(**X**=X_i)	0.20	0.55	0.25	1.00
$(x_i - \mu)^2$	$(0-117.85)^2$ =13888.62	$(137-117.85)^2$ =366.72	$(170-117.85)^2$ =2719.62	
$(x_i - \mu)^2$ P(**x**=X_i)	(13888.62)(0.2) =2777.7	(366.72)(0.55) =201.7	(2719.62)(0.25) =679.9	3659.3

So the variance of X is Var (X) $=\sigma^2=3659.3$

and standard deviation of X is $\sigma=\sqrt{3659.3}=60.49$

Mean and Variance of Linear combination of random variables

- If X and Y are random variables, then a linear combination of the random variables is

  a X +b Y   and

E(a X +b Y)=a E(X)+b E(Y)

Example 4. Leonard has invested \$6000 in Google Inc. (GooG) and \$2000 in Exxon Mobil Corp. Let X represents the change in Google's stock next month and Y represents the change in Exxon next month. We assume Google be rising 2.1% and Exxon rising 0.4% per month, How much change does Leonard expect to be in next month?

Answer: $E(6000X+2000Y)=6000E(X)+2000E(Y)$

$= 6000)(0.021)+(2000)(0.004)$

$=134$

- If X and Y are random variables, and X and Y are independent of each other, then

$$\mathbf{Var\ (a\ X + b\ Y)=a^2Var(X)+b^2Var(Y)}$$

<u>Example 5.</u> If T=4X+5Y, Var(X)=10, Var(Y)=6, and X and Y are independent, what is Var(T)?

Answer: Var(T)=Var(4X+5Y)=16Var(X)+25Var(Y)=160+150=310

<u>Chapter 2 Homework#2 (due 01/28/2016) : 2.5, 2.6, 2.21, 2.34,2.43</u>