1.1 - Basic definitions - Conditional probability

Q1

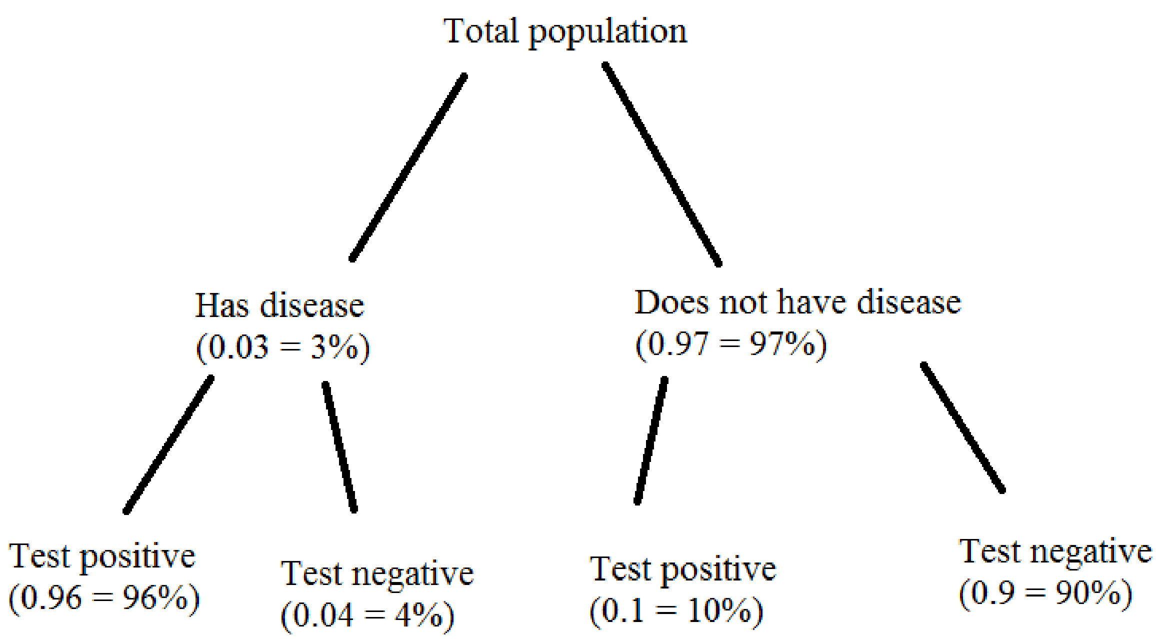
A doctor tests a patient for a certain disease. Let HD be the event that the patient Has the Disease. Let TP be the event that the Test is Positive. TP occasionally occurs when the patient is disease free; a false positive. And the complementary event TPc (tests negative) occasionally occurs when the patient has the disease; a false negative.

Most of the time, however the test does pretty well. It correctly identifies the presence of the disease in 96% of tests; P(TP|HD)=0.96. Also it correctly identifies the absence of the disease in 90% of tests; P(TPc|HDc)=0.9, where HDc means "patient does not have the disease".

The disease is present in 3 % of the population.

a. What is the probability, P(HD|TP), that a person who is tested positive for the disease, actually has the disease?

Answer: We start by drawing a tree diagram, using the data provided in the assignment:



P(TP|HD) = 0.96 P(TP^c|HD) = 0.04 P(TP|HD^c) = 0.1 P(TP^c|HD^c) = 0.9

We now use Bayes' theorem:

Applying this to the given problem (in the denominator: we add ALL the terms containing TP):
P(HD|TP) = (P(HD)*P(TP|HD))/(P(HD)*P(TP|HD) + P(TP|HD^c)*P(HD^c)) =
0.03* 0.96/(0.03*0.96 + 0.1*0.97) = 0.2289

b. What is the probability, P(HD|TP^c), that a person who is tested negative for the disease, actually has the disease?

Answer: We use the same formula as before (note that (HD^c)^c = HD):

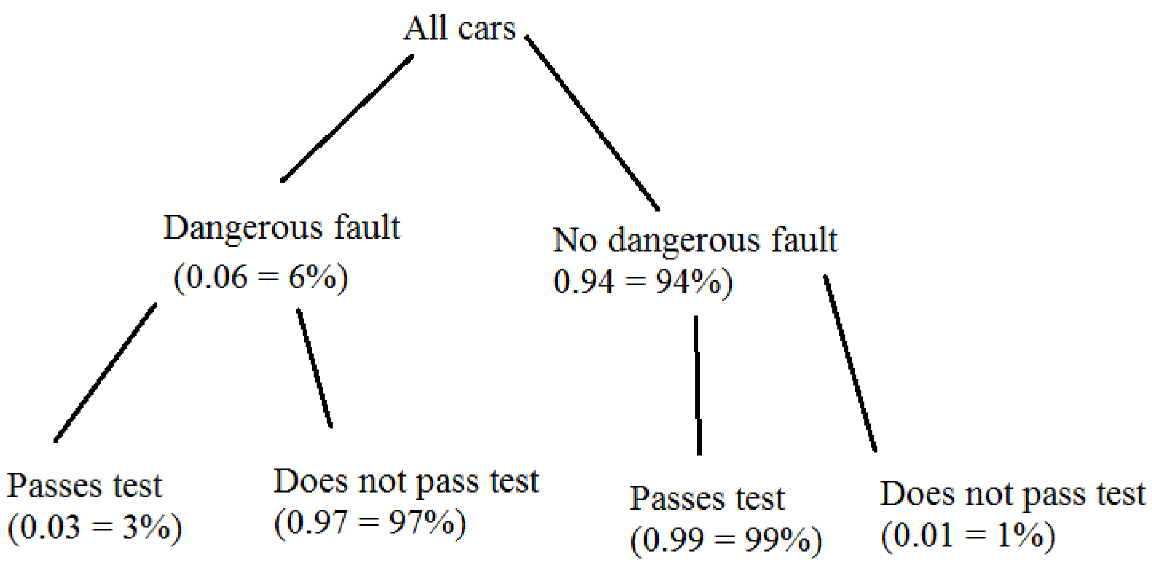
P(HD|TP^c) = (P(HD)*P(TP^c|HD))/(P(HD)*P(TP^c|HD) + P(TP^c|HD^c)*P(HD^c)) =
0.03* 0.04/(0.03*0.04 + 0.9*0.97) = 0.00137

Q2

It is known 6% of all cars over three years old suffer from at least one potentially dangerous fault. Given a car with such a fault, the government vehicle test discovers it in 97% of cases.

Also, given a fault free car, the test station passes it in 99% of cases.

What is the probability that a car that passes a test is suffering from a potentially dangerous fault?



Answer: Once again we draw a tree diagram and use Bayes' theorem:

P(Dangerous fault) = P(F) = 0.06

P(No dangerous fault) = P(F^c) = 0.94

P(Pass test) = P(P)
P(Does not pass test) = P(P^c)

P(P|F) = 0.03
P(P^c|F) = 0.97

P(P|F^c) = 0.99

P(P^c|F^c) = 0.01

P(F|P) = P(P|F)*P(F)/(P(P|F)*P(F) + P(P|F^c)*P(F^c)) =

0.03*0.06/(0.03*0.06 + 0.99*0.94) = 0.0019

Q3

Ann-Marie och Erik, arbetar på löpande bandet med att svetsa fast komponenter i moduler till TV-apparater. Ann-Marie arbetar snabbt men producerar i gengäld 4.9% felaktiga enheter. Erik, som är noggrann, gör bara fel på 0.7% av sina enheter. Under samma tid som Erik producerar 13 moduler hinner Ann-Marie med 17.

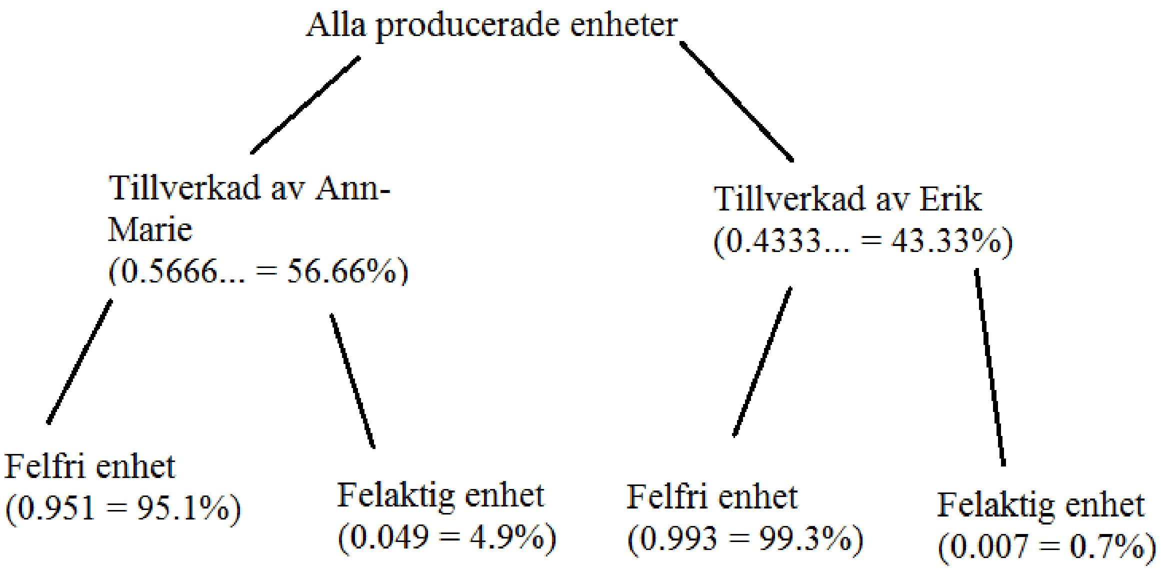
Antag att man i kvalitetskontrollen upptäcker en felaktig modul. Hur stor är chansen att denna är svetsad av Ann-Marie?

Svar: Vi ritar återigen ett träddiagram och använder Bayes' theorem:

Alla producerade enheter av Ann-Marie och Erik: 17+13 = 30
Andel producerade enheter av Ann-Marie: 17/30 = 0.5666... = P(AM) (sannolikheten att A-M har producerat enheten)

Andel producerade enheter av Erik: 13/30 = 0.4333... = P(E) (Sannolikheten att Erik har producerat den)

P(Felaktig enhet) = P(F)



P(F^c|AM) P(F|AM) P(F^c|E) P(F|E)

Notera här att vi kan byta ut E mot AM^c (komplementet till AM är ju att ”enheten INTE är tillverkad av A-M, utan av Erik”).

Vi använder oss av Bayes' theorem:

P(AM|F) = P(F|AM)*P(AM)/(P(F|AM)*P(AM) + P(F|AM^c)*P(AM^c)) =

0.049*0.5666/(0.049*0.5666 + 0.007*0.4333) = 0.9015

1.2 - Discrete random variables - Discrete random variables

Q1

Let X and Y be independent random variables with:

$E[X] = 5$ and $E[X^2] = 61$ and also $E[Y] = 9$ and $E[Y^2] = 202$.

a) Find $E[2 \cdot X + 4 \cdot Y - 9]$. 2*5 + 4*9 -9

c) Var[X] = 61 – 5^2

d) σ_X = sqrt(Var[X])

g) Find $Var[2 \cdot X + 4 \cdot Y - 9]$. 2^2*(61 - 5^2) + 4^2(202 - 9^2) i, k) 0 Specialfall

j, l) 1 Specialfall

Q2

In an experiment to graft sweet orange trees to the root of a sour orange variety, a series of five trials is conducted. Let X denote the number of grafts that fail. The density of X is given in the table below:

X	p(X)
0	0.69
1	0.12
2	0.1
3	0.05
4	0.03
5	?

a) P(5) = 1 – (Summan av alla andra p)
b) E[X] = 0*0.69 + 1*0.12 + 2*.1 och så vidare
c) Samma som b)
d) 0^2*.69 + 1^2*.12 + 2^2*.1 och så vidare
e) d) – b)^2
f) samma som e)
g) sqrt utav e)

Q3

Consider the random variable X whose density is given by
$f(x) = \frac{(x-5)^2}{C} \; for \; x = \{5,6,7\}$

a) Hitta C: f(5)+f(6)+f(7)=1
b) Coefficient framför e^(7t)? Sätt in 7 I f(x)
c) E[X] = f(5)*5+f(6)*6+F(7)*7
d) E[X^2]= f(5)*5^2+f(6)*6^2+F(7)*7^2
e) sqrt(E[X^2] – (E[X])^2)

Q4

The probability that a wildcat well will be productive is 1/15. Assume that a group is drilling wells in various parts of the country so that the status of one well has no bearing on that of any other. Let X denote the number of wells drilled to obtain the first strike.

a) Alltid 1
b) E[X] = 15 för på den 15:onde ska det ha hittats olja.
c) Find $E[X^2]$ = 15^2 * 2 – 15

d) Find σ^2 = (c) – 15)/2
e) Find σ. sqrt(d))
f) Find $P(X \geq 4)$. (14/15)^3

Q5

The probability that a child is female is 0.47, independently of the sexes of the other children.
a) In a family with 3 children, what is the probability of 1 female or fewer? 0.53^3 + (.47*0.53^2)*3
b) What is the probability of more than 1 female? 1- a)
c) probability of 1 female or more? 1-(0.53^3)

Q6

Exakt som Q6 I binomial distribution.

Q8

Av 14 komponenter är 5 felaktiga.

På hur många sätt kan man välja 5 komponenter så att man erhåller exakt 3 felaktiga om den ordning komponenterna erhålles i inte betraktas?

(5 nCr 3) * ((14-5) nCr 2)

Q13

Följande tabell visar sannolikheter för antal datorsystemfel vid ett företag under ett visst kvartal:

Antalet fel:	0	1	2	3	4+
Sannolikhet:	0.292	0.306	0.236	0.153	0.013

a) Bör bli 1.
b) färre än två fel under ett visst kvartal? P(0)+P(1)
c) fler än två fel under ett visst kvartal? P(3)+P(4)
d) åtminstone ett fel under ett kvartal? 1- P(0)
e) Kan inte beräkna väntevärde/varians. Svar: 0

Q14

Vid en kvalitetskontroll upptäcker man att de tillverkade enheterna kan ha två sorters fel, missfärgning och deformation. De två felen uppkommer oberoende av varandra. En tillverkad enhet kan alltså ha båda felen. Missfärgning finns hos 8% av alla kontrollerade enheter och deformation hos 3%. Om man vill rätta till felen så att enheten skall gå att sälja så tillkommer en kostnad på 430 kronor för enbart omfärgning. Om man vill rätta till en deformation så måste man även färga om enheten även om det inte har varit fel i färgningen från början. Tillrättning av en deformerad enhet inklusive omfärgning kostar 2000 kronor.

a) inga tilläggskostnaden ? 1 – (0.03+0.08 – 0.03*0.08)

b) tilläggskostnaden för deformation (med ommålning) ? 0.03

c) Antag att en felfri enhet kostar 5600 att tillverka från råmaterial till färdig produkt. Om man tar hänsyn till de kostnader som det kostar att rätta till de defekta enheterna, vad är det minsta pris som fabriken måste ta ut för att tillverkningen skall gå med vinst?

430*(0.08-0.03*0.08)+2000*(0.03)+5600

d) troligast antal fel: 0

1.2 - Discrete random variables - Binomial distribution

Q5

En högskola använder sig av nivågrupperad undervisning. När man tilldelar eleverna en grupp så använder man sig av resultatet från ett diagnostiskt test av typ multiple-choice. Testet består av 11 frågor med vardera 4 svarsalternativ, varav ett är rätt. För att hamna i grupp 1 så skall antal rätt besvarade frågor vara 3 eller mindre, i grupp 2 mellan 4 och 7 rätt och i grupp 3 minst 8 rätt. Antag att en student inte alls kan svaret på någon av frågorna utan väljer slumpmässigt bland svarsalternativen.

a) chansen att få alla rätt? (1/4)^11
b) chansen att få noll rätt? (3/4)^11
c) Hur många rätt kan han förvänta sig att få? (1/4)*11
d) Komma in i grupp tre (minst åtta rätt): (1/4)^8*(3/4)^3 * (11 nCr 8) + (1/4)^9*(3/4)^2 * (11 nCr 9) + (1/4)^10*(3/4)^1 * (11 nCr 10) + (1/4)^11*(3/4)^0 * (11 nCr 11)
e) Bara svarat på fyra frågor, chans för grupp två? (1/4)^4
f) Väntevärdet samma som c)
g) Vad är standardavvikelsen? Sigma = sqrt(npq) = sqrt(11*(1/4)*(1-1/4))

Q6 Man kastar ett symetrisk mynt 53 gånger. Vad är sannolikheten för att få mellan 22 och 27 kronor (inklusivt)?

CDF[BinomialDistribution[53, .5], 27] - CDF[BinomialDistribution[53, .5], (22-1)]

1.2 - Discrete random variables - Geometric distribution

Q1

It is known that the probability of being able to log on to a computer from a remote terminal is 0.7 . Let X denote the number of attempts that must be made to access the computer. Find the closed form-expression for $f(x)$, the probability density function for X. Use then this formula to calculate the first four terms of the density table:

a) f(1)=.7
b) f(2)=.7*.3
c) f(3)=.7*.3*.3

e) f(1)+f(2)+f(3)+f(4) Som mest fyra attempts

f) 1-(svaret av e) Som minst fem attempts