Module 5: Probability and Randomness
Answers to Practice exercises

PART 1: Random processes and introduction to probability

EXAMPLE 1:

a. *Relative frequency because it is based on empirical data*
b. *Personal probability*
c. *Exact (theoretical) probability based on counting the number of favorable outcomes: Pr(five cards of same suit) = (52/52)(12/51)(11/50)(10/49)(9/48) = 0.001981*
d. *Relative frequency because it is based on empirical data*
e. *Personal probability*

EXAMPLE 2: ROLL A PAIR OF DICE.

a. *(number of pairs that sum to 7)/(number of all pairs) = 6/36*
b. *(number of pairs that sum to 3)/(number of all pairs) = 2/36*
c. *(number of pairs that have a "3")/(number of all pairs) = 11/36*
d. *(number of pairs that sum to 2or 3)/(number of all pairs) = 3/36*

PART 2: Mutually exclusive events and additive rule

EXAMPLE 1: Determine which of the following events is mutually exclusive (disjoint)

a. *NOT Mutually exclusive since someone can have a high school diploma and a college degree*
b. *Mutually exclusive, since they cannot occur together.*

EXAMPLE 2: AIRLINE PASSENGERS:

a. *Yes – the probability sums up to one, and the table lists all the possible outcomes of the random variables.*
b. *Pr(2 or more passengers) = 0.057+0.009+0.001 = 0.067*

EXAMPLE 3: SOFTWARE ENGINEERS

a. *The events are not mutually exclusive since a programmer can know both Java and C++.*
b. *Pr (a programmer does not Java) = 1- Pr (programmer knows Java) = 1-0.60 = 0.40*
c. *Using the additive rule for non mutually exclusive events:*
   Pr (programmer knows either Java or C++) =
   = Pr(knows Java) + Pr(knows C++) – Pr(Knows Java and C++) =
   =0.60+0.70-0.45 = 0.85
d. *We can compute this probability using the complementary rule. Thus*
   Pr(Programmer does not know Java or C++) = 1- Pr (programmer knows either Java or C++) =
   =1-0.85 = 015
e. *Pr(Programmer knows C/C++ but not Java) = Pr(Programmer knows C/C++) – Pr(programmer knows C++ and Java) = 0.70-0.45 = 0.25*

PART 3: Independent events and multiplicative rule

EXAMPLE 1:

a. *Dependent, since by drawing a card and not replacing, we change the deck of cards*
b. *Dependent (assuming that a salary raise will affect your decision of buying a new car)*
c. *Independent (no explanation needed!)*

EXAMPLE 2:

We can assume that the event of buying an Apple is independent. We use the complementary rule and the multiplicative rule for independent events:

Pr(at least one person will buy an Apple laptop out of 10 people)

= 1-Pr(no one of the 10 people will buy an Apple laptop)=

= 1 – Pr (person will not buy an Apple laptop)^10

= 1-(0.90)^10 = 0.651

EXAMPLE 3:

a. We can assume that the three tests find the error independently

   Pr(T1 finds error, T2 finds error, T3 finds error) = Pr(T1 finds error)P(T2 finds error)P(T3 finds error) =

   *= 0.2*0.3*0.5=0.03*

b. *Pr(T1 finds error, T2 does not find error, T3 does not find error) =*

   = Pr(T1 finds error)P(T2 does not find error)P(T3 does not find error) =

   = 0.2(1-0.3)*0.5=0.07*

c. *Pr(at least one test find error)=1-Pr(no tests find error) = 1-(1-0.2)(1-0.3)(1-0.5) = 0.72*

EXAMPLE 4:

Assume that the customers act independently on each other, and use the multiplicative rule for independent events to compute the following probabilities.

a. *Pr(customer does not pay) = 0.10*
b. *Pr(two customers do not pay) = 0.10*0.10*
c. *Pr (at least one customer will pay out of 4) = 1 – P(none of the four customers will pay)=*

   *= 1 – 0.10*0.10*0.10*0.10 = 0.9999*

EXAMPLE 5: PROCESS QUALITY CONTROL (CHALLENGE!)

Use the multiplicative rule for independent events to compute the following probabilities. D indicates a defective chip and G indicates a non-defective chip.

a. Pr(4 chips are defective)=

   = Pr(“1^{st} chip =D” & “2^{nd} chip =D” & “3^{rd} chip =D” & “4^{th} chip =D”)=

   = Pr(1st chip = D) $\times$ Pr(2^{nd} chip =D) $\times$ Pr(3^{rd} chip =D) $\times$ Pr(4^{th} chip =D)=

   = 0.05 $\times$ 0.05 $\times$ 0.05 $\times$ 0.05=0.0000062

b. The probability of selecting the **first defective chip at the first try** is just the probability of selective a defective chip Pr(chip is defective) = 0.05.
c. This is different from before, because the probability of selecting the **first defective chip at the second try** means that the first chip was not defective (G), therefore

   Pr(defective chip at second try only) = Pr(first chip is not defective AND second chip is defective

= Pr(1^{st} chip =G) x Pr(2^{nd} chip=D) = 0.95 * 0.05 = 0.0475.

d. It is not the same probability computed above. The event of selecting the **fist defective chip at the fourth try** assumes that the first three chips are good. The probability is computed as

Pr(1^{st} chip = G, 2^{nd} chip= G, 3^{rd} chip=G, 4th chip=D) =

= Pr(1st chip = G) × Pr(2^{nd} chip =G) × Pr(3^{rd} chip =G) Pr(4th chip= D) =

= 0.95*0.95*0.95*0.05 = 0.043

e. This is not the same probability as there are 4 different ways in which we observe <u>one defective chip out of 4 chips</u>:

1) GGGD *(the last chip is defective)*
2) GGDG *(the next to last chip is defective)*
3) GDGG *(the third chip is defective)*
4) DGGG *(the second chip is defective)*

Each outcome has probability equal to $(.95)^3(.05)$. The probability of selecting a defective chip out of 4 chips is 4* $(.95)^3(.05)$ =0.171

f. The event of <u>selecting at least one defective chip out of 4 chips</u> is the opposite event of "all 4 chips are not defective". Thus using the complementary rule we can write:

Pr(at least a defective chip out of 4 chips) = 1 – Pr(all four chips are not defective)

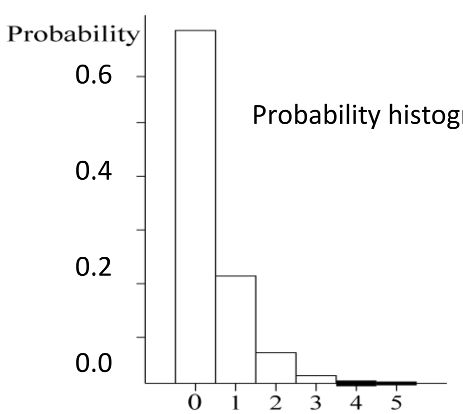
= 1 - 0.95^4=0.185.

Part 4: Random variables and probability distributions

EXAMPLE 1: REMOTE CONNECTION

a) The table displays a valid probability distribution, since all probabilities are in (0,1) and sum up to 1, and all the possible outcomes are listed.

b)



c) Pr(X=1 or X=2) = 0.7 + 0.21 = 0.91

d) E(X) = 0.7*1+0.21*2+0.063*3+0.0189*4+0.0081*5 = 1.4251. The computer connects remotely in an average of 1.43 attempts – or it takes on average between one or two attempts to connect to the remote server.

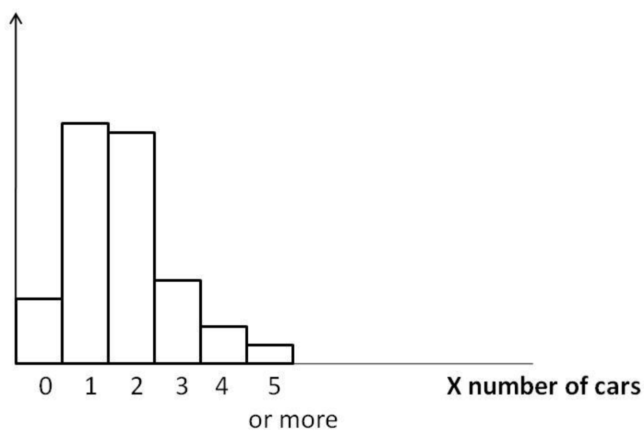
EXAMPLE 2: INSURANCE POLICIES

This is compute as the mean amount gained by the insurance (mean of X).

The mean gain per customer is E(X) = (0.10)(−$1000) + (0.90)($500) = $350. On average, the insurance company makes $350 per customer.

EXAMPLE 3: NUMBER OF CARS OF AMERICAN HOUSEHOLDS

a) The probability distribution is valid, since all the probabilities sum to one and the list of possible outcomes is complete. The probability distribution is displayed in the following histogram:



b) That's the proportion of houses with 3 cars or more, that is Pr(X>=3) = 0.13+0.05+0.02 = 0.20

c) The average number of cars per American Household is
E(X) = 0.009*0+0.36*1+0.35*2+0.13*3+0.05*4+0.02*6 = 1.77. The American households have 1.77 cars on average that is between 1 and 2 cars on average.

EXAMPLE 4: NETWORK BLACKOUTS

*Based on the probability distribution of network blackouts, the average number of blackouts is E(X) = 0*0.7 + 0.2*1+0.1*2 = 0.4. Thus the average loss due to network blackouts is $500*0.4 = $200 per day.*

EXAMPLE 5: KENO PAYOFF

The average payoff in the game of Keno is computed as E(X) = 0.25*3+0.75*(-1) = $0. If a player plays the game of Keno many times, the average win is zero dollars. The game is fair, as the expected payoff is zero dollars.

EXAMPLE 6: AMERICAN ROULETTE

a) Define X = bet payoff.

***Straight bet*:** Suppose a player bets $1 on the number 5. The probability distribution for the straight bet is as follows:

Outcome	**Loss**	**Win**

X=payoff	-$1	$35
Probability	37/38	1/38

The player wins if "5" comes up. The bet has 1 in 38 (total number of slots in the wheel) chances to win

The mean payoff (also known as house edge) is $\mu_X=35*1/38+(-1)*37/38=-2/38=-0.0526$

In the long run, players lose on average 5 cents per dollar (5% loss).

Color bet: A player bets $1 on red. The probability distribution is given below:

Outcome	**Loss**	**Win**
X=payoff	-$1	$1
Probability	20/38	18/38

The player wins if a red number comes up, and there are 18 out 38 red numbers.

The mean payoff is $\mu_X=1*18/38+(-1)*20/38=-2/38=-0.0526$

In the long run, players lose on average 5 cents per dollar (5% loss).

4-number bet: A player bets on the four numbers in the top square {1,2,4,5}.

The probability distribution is given below:

Outcome	**Loss**	**Win**
X=payoff	-$1	$8
Probability	34/38	4/38

The player wins if either one of the four numbers come up, so he/she has 4 out 38 chances to win.

The mean payoff is $\mu_X=8*4/38+(-1)*34/38=-2/38=-0.0526$

In the long run, players lose on average 5 cents per dollar (5% loss).

b) Standard deviations:

Straight bet**:** The standard deviation is $=\text{sqrt}[(35-(-0.0526))^2*1/38+(-1+0.0526)^2*37/38]=5.763$

Color bet: The standard deviation is $=\text{sqrt}[(1-(-0.0526))^2*18/38+(-1+0.0526)^2*20/38]=0.997$

Four number bet: The standard deviation is $=\text{sqrt}[(8-(-0.0526))^2*4/38+(-1+0.0526)^2*34/38]=2.7603$

The three bets have the same expected payoffs, but different standard deviations.

The straight bet has the highest standard deviation, and can be considered the riskiest bet for this game. Using the rule of thumb learned in module 1 and 2, we could say that if a player bets one dollar on the same number for many times, the average gain can most likely vary between (-11.58$, 11.47$) – computed as 2 standard deviations from the mean.