Partial Solutions for Assignment 1

(1) In a certain random experiment, let A and B be two events such that $P(A) = 0.7$, $P(B) = 0.5$, and $P((A \cup B)') = 0.1$, show that

(a) $P(A \cap B) = \underline{\ 0.30\ }$

(b) $P(A|B) = \underline{\ 0.60\ }$

(c) $P(B|A) = \underline{\ 3/7\ }$

(d) $P(A \cap B) = 0.30 \neq 0.7 \times 0.5 = P(A)P(B)$.

(e) A and B are not independent events.

(2) In a certain random experiment, let C and D be two events such that $P(C) = 0.8$, $P(D) = 0.4$, and $P((C \cup D)') = 0.12$, show that

(a) $P(C \cap D) = \underline{\ 0.32\ }$

(b) $P(D|C) = \underline{\ 0.40\ }$

(c) $P(C|D) = \underline{\ 0.80\ }$

(d) $P(C \cap D) = 0.32 = 0.8 \times 0.4 = P(C)P(D)$.

(e) C and D are independent events.

(3) The grip strengths at the end of a health dynamics course for 15 male students were

58	52	46	57	52	45	65	71	57	54	48	58	35	44	53

$[mean,\ variance,\ std,\ \sigma^2] = [\mu,\ \mathbf{s^2},\ \mathbf{std},\ \boldsymbol{\sigma^2}] = [\mathbf{53.0},\ \mathbf{78.2857},\ \mathbf{8.8479},\ \mathbf{73.0667}]$

(4) If $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.3$, then (a) $P(A \cup B) = 0.6$, (b) $P(A \cap B') = 0.1$, (c) $P(A' \cup B') = 0.7$.

(5) A typical *American roulette wheel* used in a casino has *38 slots* that are numbered $1, 2, \cdots, 35, 36, 0, 00$ respectively. The 0 and 00 are colored as *green.* Half of the remaining slots are red and half are black. Also, half of the integers between 1 and 36 inclusive are odd, half are even. A ball is rolled around the wheel and ends up in one of the 38 slots; we assume that each slot has equal probability of 1/38 and we are interested in the number of the slots in which the ball falls.

(a) The sample space $S = \{1, 2, \cdots, 35, 36, 0, 00\}$.

(b) Let $B = \{0, 00\}$, $P(B) = \frac{2}{38}$.

(c) Let $C = \{14, 15, 17, 18\}$, $P(C) = \frac{4}{38}$.

(d) Let $D = \{x \in S |\ x\ is\ odd\}$, $P(D) = \frac{18}{38}$.

(6) *See class notes* to prove that

(a) $\begin{pmatrix} n \\ k \end{pmatrix} = \begin{pmatrix} n-1 \\ k \end{pmatrix} + \begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$

(b) $\sum_{k=0}^{n} (-1)^k \begin{pmatrix} n \\ k \end{pmatrix} = 0$

(c) $\sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} = 2^n$

(7) Let an experiment be drawing five cards at random without replacement from a deck of 52 poker cards. The sample size is then $C(52,5) = 2,598,960$. Find the size of the following events.

(a) Four of a kind (four cards of equal face vaule and one card of a different value).

(key a) $C(13,1) \times C(48,1)/C(52,5) \ = \ 624/2598960 \ = \ 0.00024.$

(b) Full house (one pair and one triple of cards with equal face value).

(key b) $C(13,1) \times C(4,2) \times C(12,1) \times C(4,3)/C(52,5) \ = \ 3744/2598960 \ = \ 0.00144.$

(c) Three of a kind (three equal face values plus two cards of different values).

(key c) $C(13,1) \times C(4,3) \times C(48,2)/C(52,5) \ = \ 54912/2598960 \ = \ 0.02113.$

(d) Two pairs (two pairs of equal face value plus one card of a different value).

(key d) $C(13,2) \times C(4,2) \times C(4,2) \times C(44,1)/C(52,5) \ = \ 123552/2598960 \ = \ 0.047539.$

(e) One pair (one pair of equal face value plus three cards of different values).

(key e) $C(13,1) \times C(4,2) \times C(12,3) \times C(4,1) \times C(4,1) \times C(4,1)/C(52,5) \ = \ 1098240/2598960 \ = \ 0.42257$

(8) A box contains four marbles numbered 1 through 4. The marbles are selected one at a time without replacement. A match occurs if marble numbered k is the kth marble selected. Let the event A_i, denote a match on the ith draw, i=1,2,3,4.

(a) $P(A_i) = \frac{3!}{4!}$, for i=1,2,3,4.

(b) $P(A_i \cap A_j) = \frac{2!}{4!}$, where $1 \leq i < j \leq 4$.

(c) $P(A_i \cap A_j \cap A_k) = \frac{1}{4!}$, where $1 \leq i < j < k \leq 4$.

(d) $P(A_1 \cup A_2 \cup A_3 \cup A_4) = \frac{5}{8}$.

(e) Extend this exercise so that there are marbles in the box. Show that the probability of at least one match is

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= 1 - \tfrac{1}{2!} + \tfrac{1}{3!} - \tfrac{1}{4!} + \cdots + \tfrac{(-1)^{n+1}}{n!} \\ &= 1 - \left(1 - \tfrac{1}{1!} + \tfrac{1}{2!} - \tfrac{1}{3!} + \tfrac{1}{4!} + \cdots + \tfrac{(-1)^n}{n!}\right) \end{aligned}$$

(f) The limit of this probability as n increases without bound is $1 - e^{-1} \approx 0.6321$

(9) Consider a random experiment of casting a pair of unbiased six-sided dice and let the r.v. X equal the *smaller* of the outcomes if they are different and the common value if they are equal.

(a) The p.m.f. of r.v. X is $f(x) = \frac{13-2x}{36}$, $x = 1, 2, 3, 4, 5, 6$.

(b) `X=1:6;  Y=(13-2*X)/36;  bar(X,Y,0.8);  title('y=f(x)=(13-2x)/36')`

(c) $E(X) = \frac{91}{36}$ and $Var(X) = \frac{2555}{1296}$.

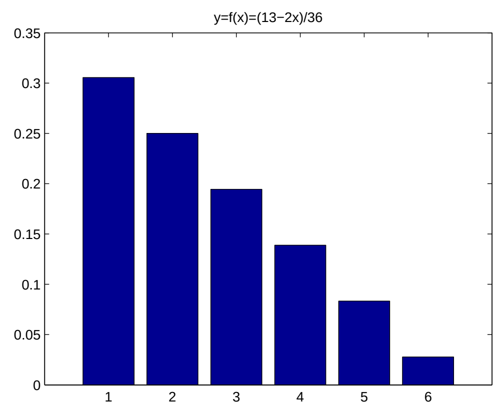

Figure 1: 9(b) Histogram of Probability.

(10) The number of defects on a printed circuit board is a random variable (r.v.) X with the probability mass function (p.m.f.) given by

$$P(X = i) = \gamma/(i+1), \quad i = 0, 1, 2, 3, 4$$

(a) The constant $\gamma = \frac{60}{137}$.

(b) The mean of X is $E(X) = \frac{163}{137}$.

(c) The variance of X is $Var(X) = \frac{33300}{18769}$.

```
% h1p10.m - Solution for Problem 10
%
format rat;
p=[1, 1/2, 1/3, 1/4, 1/5];
r=1/(sum(p));   % r=60/137;
f=[r/1, r/2, r/3, r/4, r/5];
X=0:1:4;
EX=X*f';
VarX=((X-EX).^2)*f';
[r, EX, VarX]   % VarX is an approximate rational to 33300/18769
format short
[r, EX, VarX]   % Check the precision of the numerical solution
```

(11) Bean seeds from supplier A have 85% germination rate and those from supplier B have a 75% germination rate. A seed packaging company purchases 40% of their bean seeds from supplier A and 60% from supplier B and mixes these seeds together.

(a) The probability that a seed selected at random from the mixed seeds will germinate is $P(G) = P(G|A)P(A) + P(G|B)P(B) = 0.85 \times 0.4 + 0.75 \times 0.6 = 0.79$.

(b) Given that a seed germinates, the probability that the seed was purchased from supplier A is $P(A|G) = \frac{P(A \cap G)}{P(G)} = \frac{P(G|A)P(A)}{P(G)} = \frac{34}{79}$.

(c) Given that a seed germinates, find the probability that the seed was purchased from supplier B is $P(B|G) = \frac{P(B \cap G)}{P(G)} = \frac{P(G|B)P(B)}{P(G)} = \frac{45}{79}$.

(12) Some ornithologists were interested in the clutch size of the common gallinule. They observed the number of eggs in each of 117 nests which are listed below.

7	5	13	7	7	8	9	9	9	8	8	9	9	7	7
5	9	7	7	4	9	8	8	10	9	7	8	8	8	7
9	7	7	10	8	7	9	7	10	8	9	7	11	10	9
9	4	8	6	8	9	9	9	8	8	5	8	8	9	9
14	10	8	9	9	9	8	7	9	7	9	10	10	7	6
11	7	7	6	9	7	7	6	8	9	4	6	9	8	9
7	9	9	9	9	8	8	8	9	9	9	8	10	9	9
8	5	7	8	7	6	7	7	7	6	5	9			

(a) Construct a frequency table for these data.

(b) Draw a histogram.

(c) What is the *mode* (the typical clutch size)?

```
% h1p12.m - The number of eggs in 117 nests of the common gallinule
%
N=14; m=117;
fin=fopen('dataP12.txt','r');
fout=fopen('outP12.txt','w');
fgetl(fin);
H=fscanf(fin,'%d',m);  fclose(fin);
X=1:1:14;  Y=zeros(1,14);
for j=1:m   k=H(j);  Y(k)=Y(k)+1;  end
fprintf(fout,'Eggs  Freq\n')
for i=1:N  fprintf(fout,'%4d   %4d\n',X(i),Y(i));  end
fclose(fout);
bar(X,Y,0.8)
legend('mode=9, [3,5,7,27,26,37,8,2,0,1,1]',2)
title('Solution for Problem 12')
```

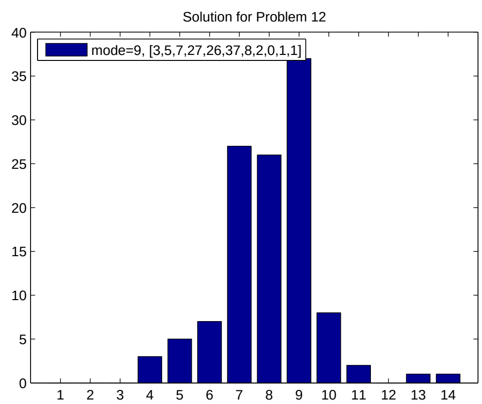

Figure 2: (12) Frequency of the number of eggs.

(13) The following data give the ACT *math* and ACT *verbal* scores, say, (x, y), for 15 students:

(16,19)	(18,17)	(22,18)	(20,23)	(17,20)
(25,21)	(21,24)	(23,18)	(24,18)	(31,25)
(27,29)	(28,24)	(30,24)	(27,23)	(28,24)

(a) Verify that $\bar{x} = 23.8$, $\bar{y} = 21.8$, $s_x^2 = 22.457$, $s_y^2 = 11.600$, $r = 0.626$.

(b) Find the equation of the best fitting line.

(c) Plot the 15 points and the line on the same graph.

```
%
% Script file: plotACT.m - Problem 13 of H1
% A Least Squares Line Fit for Data on ACT(math, verbal) test
%
X=[16, 18, 22, 20, 17, 25, 21, 23, 24, 31, 27, 28, 30, 27, 28]; % math
Y=[19, 17, 18, 23, 20, 21, 24, 18, 18, 25, 29, 24, 24, 23, 24]; % verbal
ux=mean(X); uy=mean(Y);
sx2=var(X); sy2=var(Y);
r=(X-ux)*(Y-uy)'/14/(sqrt(sx2*sy2));           % correlation coefficient
P=polyfit(X,Y,1);                               % P=polyfit(X,Y,3);
[ux,uy,sx2,sy2,r]
Yh=P(1)*X + P(2);                     % Yh=P(1)*X.^3+P(2)*X.^2+P(3)*X+P(4);
top=norm(Yh-uy,2);
bot=norm(Y-uy,2);
R2=(top*top)/(bot*bot);
plot(X,Y,'bo',X,Yh,'r-');
legend('R^2 Statistics = 0.3916','\itY=0.4497\itX+11.0961',4);
title('Best Line Fitting on Data Set ACT (math,verbal)')
```

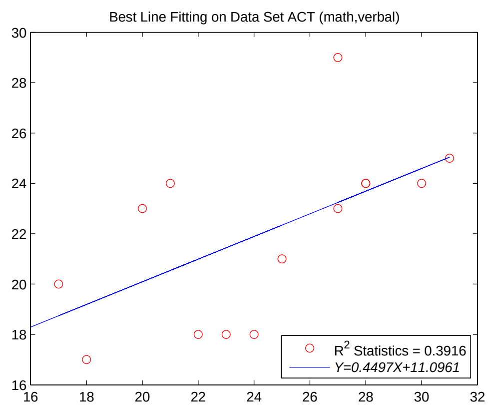

Figure 3: (13) Best Line Fitting on Data Set ACT (math verbal).

(14) The final scores of *59* students taking CS3332 course in Fall, 1999 are listed below.

61	72	77	58	67	70	76	70	76	83
42	58	49	74	65	55	90	80	31	61
53	82	90	51	48	55	84	70	48	76
61	76	70	70	66	50	80	73	77	43
71	99	66	63	63	52	54	80	67	29
52	83	62	60	61	86	61	70	73	

(a) Find the order statistics of this set of data.

(b) Find the 25th, 75th percentiles, and the median.

(c) Find five-number summary of these data.

(d) Draw a box-and-whisker diagram.

```
% boxp14.m - Read Input Data from CS3332 (1999) Scores
%
T=[61,72,77,58,67, 70,76,70,76,83, 42,58,49,74,65, 55,90,80,31,61,...
53,82,90,51,48, 55,84,70,48,76, 61,76,70,70,66, 50,80,73,77,43,...
71,99,66,63,63, 52,54,80,67,29, 52,83,62,60,61, 86,61,70,73];
S=sort(T,'ascend');
% S=[29,31,42,43, 48,48,49,50, 51,52,52,53, 54,55,55,58, 58,60,61,61,
%    61,61,61,62, 63,63,65,66, 66,67,67,70, 70,70,70,70, 70,71,72,73,
%    73,74,76,76, 76,76,77,77, 80,80,80,82, 83,83,84,86, 90,90,99];
boxplot(T,'orientation','horizontal');
text(29,0.7,'(c) [29, 55, 67, 76, 99]')
title('box-whisker plot')
```

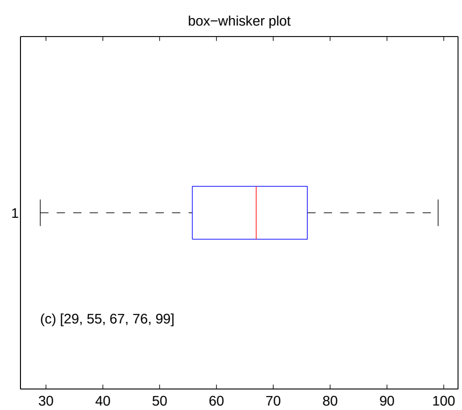

Figure 4: (14) A Box-Whisker Plot for 59 Scores in 1999.