0.1 Probability Questions

0.1.1 Basic Probability Questions

Suppose a pair of fair dice is thrown.

(a) What is the probability of getting a sum of 9 from two throws of a dice

Find the probability that the sum is 10 or greater if

(b) a 5 appears on the first die,

(c) a 5 appears on at least one of the dice.

0.1.2 Basic Probability Questions

Tickets numbered 1 to 20 are mixed up and then a ticket is drawn at random. What is the probability that the ticket drawn has a number which is a multiple of 3 or 5

0.1.3 Exercise : Probability of Two Dice Rolls

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

The table below gives the number of thunderstorms reported in a particular summer month by 100 meteorological stations.

Number of thunderstorms:	0	1	2	3	4	5
Number of stations:	22	37	20	13	6	2

(a) Calculate the sample mean number of thunderstorms.

(b) Calculate the sample median number of thunderstorms.

(c) Comment briefly on the comparison of the mean and the median.

In a certain large population 45% of people have blood group A. A random sample of 300 individuals is chosen from this population. Calculate an approximate value for the probability that more than 115 of the sample have blood group A.

If X is the number in the sample with group A, then X has a binomial (300, 0.45) distribution, so

$$E[X] = 300 \times 0.45 = 135$$

and

$$Var[X] = 300 \times 0.45 \times 0.55 = 74.25$$

. Then, using the continuity correction,

$$P(X > 115) = P(X > 115.5)$$
$$1 - \frac{115.5 - 135}{\sqrt{74.25}}$$

The mean of a sample of 30 claims is \$5,200. Six have mean of \$ 8000 (i.e. group 1) Ten have mean of \$ 3100 (i.e. group 2)

Compute the mean for the remaining claims

$$\text{Total Costs} = (CostforGroup1) + (CostforGroup2) + (CostfromGroup3)$$

- Total Cost for all three groups : $\$5200 \times 30 = \$156000$
- Cost for Group 1 : $\$8000 \times 6 = \$48000$
- Cost for Group 2 : $\$3100 \times 10 = \$31000$

Necessarily the cost for group 3 is \$77000

The mean claim for group 3 is therefore

$$\frac{\$77000}{14} = \$5500$$

Poisson/Binomial/Exponential

- Poisson

  Find P(X=0) for Poisson Mean (m=0.5)

$$P(X = 0) = \frac{e^{-0.5}}{0!} = 0.606$$

- Binomial
- Exponential

  No Claim in the next two years

$$= (0.606)^2 = 0.368$$

- Time Until Next Claim

  $\mu = 0.5$

  $T \approx exp(0.5)$

- $P(XT > 2) = exp(-1) = 0.368$

Consider Two events A, B such that

- P(A) = 0.3
- P(A $\cap B$) = 0.1

Find the minimum possible value of $P(A|B)$

$$P(A|B) = \frac{P(A \cup B)}{P(B)}$$

- $P(A \cup B) Maximum = 1$
- $P(A \cup B) Minimum = 0.3$

Minimum

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$0.3 = 0.3 + P(B) - 0.1$$
$$P(B) = 0.1$$

Maximum

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$1 = 0.3 + P(B) - 0.1$$
$$P(B) = 0.8$$

$$P(A|B)_{MIN} = \frac{P(A \cup B)}{P(B)} = \frac{0.1}{0.8} = 0.125$$
$$P(A|B)_{MAX} = \frac{P(A \cup B)}{P(B)} = \frac{0.1}{0.1} = 1$$

0.1.4 Graphical Methods - Question 3

The masses of 30 human males and 30 arabian stallions were observed. Their masses (in lbs) are given below

```
Humans
106, 120, 130, 138, 145, 151, 156, 161, 166, 171
176, 180, 185, 189, 194, 198, 203, 208, 212, 217
223, 228, 234, 240, 247, 255, 264, 276, 290, 313
```

```
Stallions
808, 824, 835, 843, 851, 857, 862, 868, 872, 877
881, 886, 890, 894, 898, 902, 906, 910, 914, 919
923, 928, 932, 938, 943, 949, 957, 965, 976, 992
```

a) Draw histograms for these samples and compare them with respect to shape, centrality and relative dispersion.

b) Calculate the medians of these samples (from the raw data).

0.1.5 Descriptive Statistics - Question 4

The following data give the marks of 10 students in a test (out of 20 marks). Calculate

i) the median

ii) the mean

iii) the range

iv) the standard deviation

v) The Inter-Quartile Range

$$12, 17, 7, 11, 18, 6, 14, 15, 11, 9.$$

0.1.6 Conditional Probability - VaVa-Voom

VaVaVoom The plant in Austria produces 80% of the cars. The plant in Belgium produces 20% of the cars. A randomly chosen car was build at Austrian plant A randomly chosen car was built at the Belgian plant S: A randomly chosen car has standard The mean and standard deviation of the following

We are told the following piece of information $\bar{x} = 44$ So what is the coefficient of determination?

The duration, in months, of the construction phase of a number of motorway projects were collected and tabulated as follows Calculate the mean value of the durations. Calculate the variance of the data set. Calculate the standard deviation. Calculate the coefficient of variation. Firstly, what is the sample size?

$$cv = \frac{s}{\bar{x}} \times 100\%$$

$$\bar{x} = \frac{49 + 55 + 43 + 45 + 41 + 33 + 42}{7}$$

$$\bar{x} = 44$$

$$s^2 = 47$$

$$s = \sqrt{47}$$

$$\frac{(49-44)^2 + (55-44)^2 + (43-44)^2 + (45-44)^2 + (41-44)^2 + (33-44)^2 + (42-44)^2}{7-1}$$

$$25 + 121 + 1 + 1 + 9 + 121 + 4 = 282/6 = 47 s = \sqrt{47}$$

$= 15.58\%$

0.1.7 Binomial Distribution

According to a recent poll, approximately seventy percent of U.S. adults drink alcohol. Suppose 5 U.S. adults are randomly selected. Let represent the number of adults in the sample who drink alcohol. Use the binomial probability formula, the binomial probability table, or your calculator to find the following probabilities.

- a. That exactly 2 adults in the sample drink alcohol. = 0.1323
- b. That at least three adults in the sample drink alcohol. = P(3) + P(4) + P(5) = 0.3087 + 0.36015 + 0.16807 = 0.83692
- Alternatively, you can use binomcdf on the calculator: P(at least 3) = 1 P(2 or fewer) = 1 binomcdf(5, 0.70, 2) = 0.83692
- c. That everyone in the sample drinks alcohol. = 0.16807

0.2 Discrete Probability Distributions

0.2.1 Confidence Intervals for Proportions

- In a survey conducted by a mail order company a random sample of 200 customers yielded 172 who indicated that they were highly satisfied with the delivery time of their orders.
- Calculate an approximate 95% confidence interval for the proportion of the company's customers who are highly satisfied with delivery times.

$$p = \frac{172}{200} = 86\%$$

$$\frac{p(100-p)}{n} = \frac{86 \times 14}{200}$$

0.2.2 Exercise: Counting

Given S is the set of all 5 digit binary strings, E is the set of a 5 digit binary strings beginning with a 1 and F is the set of all 5 digit binary strings ending with two zeroes.

(a) Find the cardinality of S, E and F.

(b) Draw a Venn diagram to show the relationship between the sets S, E and F. Show the relevant number of elements in each region of your diagram.

0.2.3 Exercise: Combinations

A committee of 4 must be chosen from 3 females and 4 males.

- In how many ways can the committee be chosen.
- In how many cans 2 males and 2 females be chosen.
- Compute the probability of a committee of 2 males and 2 females are chosen.
- Compute the probability of at least two females.

Part 1
We need to choose 4 people from 7:
This can be done in

$${}^{7}C_4 = \frac{7!}{4! \times 3!} = \frac{7 \times 6 \times 5 \times 4!}{4! \times 3!} = 35 \text{ ways.}$$

Part 2
With 4 men to choose from, 2 men can be selected in

$${}^{4}C_2 = \frac{4!}{2! \times 2!} = \frac{4 \times 3 \times 2!}{2! \times 2!} = 6 \text{ ways.}$$

Similarly 2 women can be selected from 3 in

$$^3C_2 = \frac{3!}{2! \times 1!} = \frac{3 \times 2!}{2! \times 1!} = 3 \text{ ways.}$$

Part 2
Thus a committee of 2 men and 2 women can be selected in $6 \times 3 = 18$ ways.

Part 3

The probability of two men and two women on a committee is

$$\frac{\text{Number of ways of selecting 2 men and 2 women}}{\text{Number of ways of selecting 4 from 7}} = \frac{18}{35}$$

Part 4

- The probability of at least two females is the probability of 2 females or 3 females being selected.
- We can use the addition rule, noting that these are two mutually exclusive events.
- From before we know that probability of 2 females being selected is 18/35.
- We have to compute the number of ways of selecting 1 male from 4 (4 ways) and the number of ways of selecting three females from 2 (only 1 way)
- The probability of selecting three females is therefore $\frac{4\times1}{35} = 4/35$
- So using the addition rule

$$Pr(\text{ at least 2 females }) = Pr(\text{ 2 females }) + Pr(\text{ 3 females })$$

$$Pr(\text{ at least 2 females }) = 18/35 + 4/35 = 22/35$$

0.2.4 Worked Example 1

An ordered sequence of four digits is formed by choosing digits without repetition from the set $\{1, 2, 3, 4, 5, 6, 7\}$
.

(i) the total number of such sequences; (780)

(ii) the number of sequences which begin with an odd number; (480) N(A)

(iii) the number of sequences which end with an odd number; (480) (NB)

(iv) the number of sequences which begin and end with an odd number;(240)

(v) the number of sequneces which begin with an odd number or end with an odd number or both; (720)

(vi) the number of sequences which begin with an odd number or end with an odd number but not both. (480)

Question 15. (6 marks)

Orders for a computer are summarized by the number of optional features that are requested as follows:

Number of optional features X	0	1	2	3	4
Probability P(X)	.35	.25	.20	.10	.10

a. Calculate the expected number of optional features and the variance.

0.2.5 HT for Difference of Props

A survey, carried out at a major flower and gardening show, was concerned with the association between the intention to return to the show next year and the purchase of

goods at this year s show. There were 220 people interviewed and of these 101 had made a purchase; 69 of these people said they intended to return next year. Of the

119 who had not made a purchase, 68 said they intended to return next year.

By testing the difference between the proportions of purchasers and non-purchasers who intend to return next year, examine whether there is sufficient

evidence to justify concluding that the intention to return depends on whether or not a purchase was made.

H0: population proportions of those who intend to return are equal H1: population proportions of those who intend to return are NOT equal

Proportion of purchasers 1 69 /101; proportion of non-purchasers 2 68 /119

Observed value of D = 0.1117

Estimated standard error of D = 6.558%

0.3 Poisson Approximation

n = 25 p = 0.1, 0.2

Poisson approximation of Binomial (letting $m = np$

- $m_1 = 2.5$
- $m_2 = 5$

Find $P(X \geq 5)$

$P(X \leq 4) = 1 - P(X \geq 5)$

From Tables 0.89118 0.44049

(Rest : Compare to Real Answers)

A random sample of 10 is taken from a normal distribution of $\mu = 20$ and $\sigma^2 = 1$. Let s^2 be the sample variance.

Find $P(S^2 > 1)$

Solution

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$9S^2 \sim \chi^2_9$$

$$P(S^2 > 1) = P(\chi_9^2 > 9) = 1.05627 = 0.437$$

PMS Spring 2008

$$f(x, y) = \frac{4}{3}(1 - xy) \; 0 < x < 1, 0 < y < 1$$

The Marginal PDF of X and Y is given by

$$f(x) = \frac{2}{3}(2 - x) \; 0 < x < 1$$

$$f(x, y) = \frac{2}{3}(2 - y) \; 0 < y < 1$$

Show that the conditional expectation of Y given X is given by

$$f(y|x) = \frac{2(1 - xy}{2 - 7} \; 0 < y < 1$$

Solution

$$f(y|x) = \frac{f(x, y)}{f(x)} = \frac{\frac{4}{3}(1 - xy)}{\frac{2}{3}(2 - x)} = \frac{2(1 - xy)}{2 - x}$$