1. Given a normally distributed population having a mean length of $\mu = 5$ cm and $\sigma = 2$, what is the probability of random sampling an individual whose length (x) is greater than 4 but less than 7 ? (3 marks)

Here $\mu = 5$, $\sigma = 2$, $z = (x - \mu)/\sigma$

$Pr(x < 4) = Pr(z < -0.5) = Pr(z > 0.5) = 0.30854$

$Pr(x > 7) = Pr(z > 1.0) = 0.15866$

The desired probability is given by 1 - 0.15866 - 0.30854 = 0.5328

Probability of sampling an individual in the specified range is 0.5329

2. Given a normally distributed population with $\mu = 5$ and $\sigma = 12$, what is the probability of obtaining a random sample of n = 9 individuals having a mean $\bar{x} > 4$? (3 marks)

$\sigma_{\bar{x}} = \sigma/\sqrt{n}$, so $\sigma_{\bar{x}} = 12/\sqrt{9}$, $\sigma_{\bar{x}} = 4.0$ ½ and $z = (\bar{x} - \mu)/\sigma_{\bar{x}}$

$Pr(\bar{x} > 4) = 1 - Pr(\bar{x} < 4) = 1 - Pr(z < -0.25) = 1 - Pr(z < 0.25)$

$= 1 - 0.40129$
$= 0.59871$

So the probability of a mean greater than 4 is 0.59871

3) At one particular farm, the mean weight of a strawberry was predicted to be 30 grams. You test this hypothesis by randomly sampling and weighing 6 strawberries. (5 marks)

Weight of strawberries (in grams): **29, 32, 27, 24, 30, 26**
Use the most appropriate hypothesis test to address this question.

Ho: $\mu = 30$

Ha: $\mu \neq 30$

Include calculations in here

$\bar{x} = 28.0$ $s = 2.898$ $n = 6$

SE = 1.183

$t = (28.0 - 30)/1.183$

$t = -1.69$

DF = 6 -1 =5

$t_{crit,df=5,\alpha 2=0.05} = -2.57$

State assumptions of the hypothesis test: strawberry weights are normally distributed

Decision made about the null hypothesis and a short (1 sentence) statement of conclusions.

Don't Reject Ho: because $t=-1.69 > t_{crit} = -2.57$

Conclude that there is no evidence that strawberry weights differ from 30 grams

4) The ratio of fruitfly eye colours from a particular cross is expected to be:

9/16 Red : 3/16 Vermilion : 3/16 Purple : 1/16 White.

A geneticist obtains the following numbers in a random sample of progeny:

30 Red , 15 Vermilion, 10 Purple, 5 White

Conduct the most appropriate hypothesis test. (6 marks)

Ho: ratio of eye colours is 9/16 Red : 3/16 Vermilion : 3/16 Purple : 1/16 White

Ha: ratio of eye colours is NOT 9/16 Red : 3/16 Vermilion : 3/16 Purple : 1/16 White

Include calculations in here

Sample size n = 60

	red	vermilion	purple	white
observed	30	15	10	5
expect	9/16 x 60 = 33.75	3/16x60= 11.25	3/16x60= 11.25	1/16x60= 3.75

Since the blue "expected" is less than 5, we'll pool the purple and white then calc chisquared

Pooled table

	red	vermilion	Purple and white
observed	30	15	15
Expect	33.75	11.25	15.0

$X^2_{calc} = (30-33.75)^2/33.75+(15-11.25)^2/11.25 +(15-15.0)^2/15.0$

$X^2_{calc} = 1.67$

DF = 3 -1 = 2

$\chi_2^2 = 5.99$

Decision made about the null hypothesis and a short (1 sentence) statement of conclusions.

We don't reject Ho because $X^2_{test} = 1.67 > \chi_2^2 = 5.99$.

There is no evidence for a departure from the expected ratios .

5) It is hypothesized that the proportion of mutant white squirrels in Exeter Ontario is 0.4 (or 40 %), the remainder are grey in colour. A biologist believes that this is an overestimate of the true proportion of white mutants. To test the hypothesis, they obtain a random sample with the following counts:

13 Grey and 2 White. (5 marks)

Conduct the most appropriate hypothesis test.

Ho: proportion white squirrels, $p = 0.4$

Ha: $p < 0.4$

Include calculations in here

Estimated proportion of white is: $\hat{p}=2/15 = 0.133$,   n =15
Binomial dist
$Pr (x) = n!/(x!(n-x)!)\ p^x\ (1-p)^{n-x}$

Calc prob of result as extreme or more extreme (so 2, 1, 0 white squirrels)

$Pr (x=0) = 15!/(0!15!)\ .4^0\ .6^{15} =$ 0.00047

$Pr (x=1) = 15!/(1!14!)\ .4^1\ .6^{14} =$ 0.0047

$Pr (x=2) = 15!/(2!13!)\ .4^2\ .6^{13} =$ 0.02194

Sum = 0.027114

Therefore Pvalue = 0.027

Decision made about the null hypothesis and a short (1 sentence) statement of conclusions.

Reject Ho because Pvalue of $0.027 < 0.05$.

The frequency of white squirrels is less than 0.4 and is approximately 0.13

6) An ecologist wishes to determine if herons and egrets prefer to eat the same kinds of amphibians. You observe a random sample of these birds recording for each one the first kind of food it consumes. (5 marks)

Conduct the most appropriate statistical test

	Skink	*Salamander*	*Frog*
Heron	40	5	55
Egret	60	25	15

Ho: Food type and bird species are independent

Ha: Food type and bird species are NOT independent

Include calculations in here
Observed table

	Skink	*Salamander*	*Frog*
Heron	40	5	55
Egret	60	25	15

Expected table

	Skink	*Salamander*	*Frog*
Heron	50	15	35
Egret	50	15	35

$X^2_{calc} = \sum(\text{obs-exp})^2 / \text{exp} = (40-50)^2/50 + \ldots + (15-35)^2/35$

$X^2_{calc} = 40.19$

DF =2

$\chi_2^2 = 5.99$

Decision made about the null hypothesis and a short (1 sentence) statement of conclusions.
Reject Ho because $X^2_{calc} = 40.19 > \chi_2^2 = 5.99$.
Observing tables we see too many frogs eaten by herons compared to expected, and too few frogs eaten by Egrets which seem to prefer skinks and salamanders

7) A scatologist wishes to determine if rat droppings (i.e. faecal pellets) are randomly distributed in the basement of the Farkooharsen building at Yokem University. He randomly samples 200 equal-sized floor tiles in the basement and has his PhD student count the number of rat droppings on each. Test the hypothesis above.
(7 marks)

Results: 100 tiles had 0 droppings; 80 had 1 dropping; 20 had 2 droppings

Ho: droppings are randomly distributed in basement

Ha: droppings are NOT randomly distributed in basement

Include calculations in here

$Pr(x) = e^{-\mu} \mu^x / x!$, we must estimate the mean, $\bar{x}$, and the standard deviation, s. n = 200

$\bar{x} = (0 \times 100 + 1 \times 80 + 2 \times 20)/200$, $\bar{x} = 0.6$ droppings per tile

$\sum x^2 = (0^2 \times 100 + 1^2 \times 80 + 2^2 \times 20) = 160$

So $s^2 = (110 - 70^2/200)/199$, or $s^2 = 0.442$

Num starfish	Obs num	Poisson probs	Expect number
0	100	$e^{-.6}\, 0.6^0 / 0! = 0.5488$	109.76
1	80	$e^{-.6}\, 0.6^1 / 1! = 0.3292$	65.86
2	20	= 0.1219	24.38
		Get last one by subtracting sum from 1	

$X^2_{calc} = \sum(\text{obs-exp})^2/\text{exp} = (100-109.76)^2/109.76 + \ldots + (20-24.38)^2/24.38$

$X^2_{calc} = 4.69$

df = 3-1-1 =1

critical value of with 1 degrees of freedom, $\chi_1^2 = 3.84$

Decision made about the null hypothesis and a short (1 sentence) statement of conclusions.
Reject null hypothesis since calculated $X^2_{test=}$ 4.69 is greater than the critical value $\chi_1^2 = 3.84$ so the droppings are not randomly distributed
The fact that the variance $s^2 = 0.44$ is greater than the mean, $\bar{x} = 0.6$, indicates that the droppings have a more uniform (or dispersed) distribution.

8) Write all the SAS programming statements necessary to carry out the statistical analysis below. Include the data as well. (6 marks)
A geneticist expects the following proportions of flower colours in a population of chickory plants: 0.5 Purple: 0.3 Pink : 0.2 White

In a random sample they obtain 40 Purple, 33 Pink, 12 White

```
DATA FLOWERS;
 INPUT COLOUR $ NUMB;
 DATALINES;
 Purple 40
 Pink   33
 White 12
 ;
 PROC FREQ ORDER=DATA;
   WEIGHT NUMB;
   TABLES COLOUR/CHISQ NOCUM TESTP=(0.5 0.3 0.2);
 RUN;
```