Question 1

Your answer is CORRECT.

Hospital records show that 12% of all patients are admitted for heart disease, 28% are admitted for cancer (oncology) treatment, and 6% receive both coronary and oncology care. What is the probability that a randomly selected patient is admitted for something other than coronary care? (Note that heart disease is a coronary care issue.)

a) 0.72

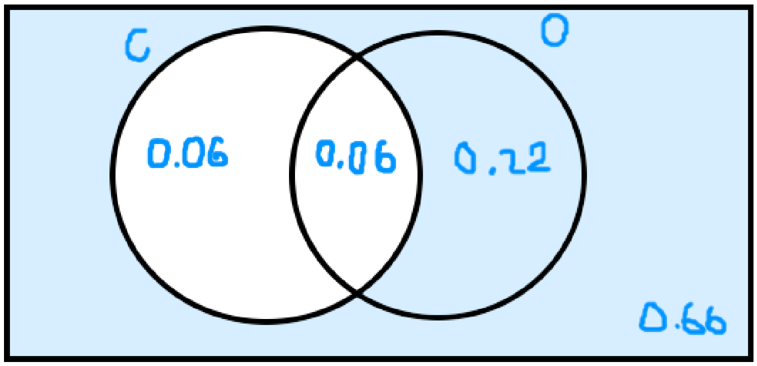
b) 0.94

c) 0.82

d) 0.66

e) 0.88

f) None of the above.




```
> 0.22+0.66
[1] 0.88
>
```

Question 2

Your answer is CORRECT.

The probability that a randomly selected person has high blood pressure (the event H) is P(H) = 0.4 and the probability that a randomly selected person is a runner (the event R) is P(R) = 0.5. The probability that a randomly selected person has high blood pressure and is a runner is 0.2. Find the probability that a randomly selected person has high blood pressure, given that he is a runner.

a) 1

b) 0.50

c) 0

d) 0.22

e) 0.40

f) None of the above.

$P(H)=0.4$
$P(R)=0.5$
$P(H\cap R)=0.2$

$$P(H|R)=\frac{P(H\cap R)}{P(R)}$$

```
> 0.2/0.5
[1] 0.4
>
```

Question 3

Your answer is CORRECT.

Given the first type of plot indicated in each pair, which of the second plots could not always be generated from it?

a) stem and leaf, histogram

b) histogram, box plot

c) dot plot, histogram

d) stem and leaf, dot plot

Question 4

Your answer is CORRECT.

Given a data set consisting of 33 unique whole number observations, its five-number summary is:

$$[14, 27, 39, 50, 65]$$

min, Q_1, Q_2, Q_3, max

How many observations are less than 39?

a) 17

b) 16

c) 38

d) 15

percentile to rank formula: $nP + 0.5$ (where n is number of data measurements, and P is percentile)

39 is Q2, the 50th percentile.

```
> 33*0.50 + 0.5
[1] 17
> # less than 39, so last one not counted (i think lol)
> 17 - 1
[1] 16
>
```

Question 5

Your answer is CORRECT.

Given the following sampling distribution:

X	-17	-9	-5	8	13
P(X)	$\frac{1}{50}$	$\frac{9}{100}$	$\frac{1}{20}$	$\frac{9}{100}$	___

What is the mean of this sampling distribution?

a) 9.1

b) 8.8

c) −0.3

d) −2.0

e) 9.2

f) None of the above

$$E[X] = \sum x \cdot p(x)$$

```
> 1 - (1/50 + 9/100 + 1/20 + 9/100)
[1] 0.75
> x = c(-17, -9, -5, 8, 13)
> px = c(1/50, 9/100, 1/20, 9/100, 0.75)
> sum(x*px)
[1] 9.07
>
```

Question 6

Your answer is CORRECT.

In testing a certain kind of missile, target accuracy is measured by the average distance X (from the target) at which the missile explodes. The distance X is measured in miles and the sampling distribution of X is given by:

X	0	10	50	100
P(X)	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{10}$

Calculate the variance of this sampling distribution.

a) 438.7

b) 36.3

c) 2258.7

d) 1316.0

e) 52.0

f) None of the above

$$E[X^2] - (E[X])^2$$

```
> x = c(0, 10, 50, 100)
> px = c(1/10, 1/5, 2/5, 3/10)
> sum((x^2)*px) - (sum(x*px))^2
[1] 1316
>
```

Question 7

Your answer is CORRECT.

In testing a new drug, researchers found that 5% of all patients using it will have a mild side effect. A random sample of 7 patients using the drug is selected. Find the probability that exactly two will have this mild side effect.

a) 0.05062

b) 0.07062

```
> dbinom(2, 7, 0.05)
[1] 0.0406235
>
```

c) 0.04062

d) 0.06062

e) 0.09062

f) None of the above

Question 8

Your answer is CORRECT.

Let X be the amount of time (in hours) the wait is to get a table at a restaurant. Suppose the cdf is represented by

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^3}{64} & 0 \leq x \leq 4 \\ 1 & x > 4 \end{cases}$$

Use the cdf to determine P(X ≥ 2).

a) 0.1250

$$P(X \geq 2) = 1 - P(X \leq 2) = 1 - F(2)$$

b) 0.8750

```
> 1 - ((2^3)/64)
[1] 0.875
>
```

c) 0.4694

d) 0.9387

e) 0.1184

f) None of the above

Question 9

Your answer is CORRECT.

Let X be the amount of time (in hours) the wait is to get a table at a restaurant. Suppose the cdf is represented by

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2}{25} & 0 \leq x \leq 5 \\ 1 & x > 5 \end{cases}$$

Use the cdf to determine E[X].

a) 6.67

$$f(x) = \frac{d}{dx}[F(x)] = \begin{cases} \dfrac{2x}{25} & 0 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

b) 3.00

c) 3.33

$$E[x] = \int_0^5 x f(x)\,dx = \int_0^5 \frac{2x^2}{25}\,dx = \left.\frac{2x^3}{75}\right|_0^5 =$$

```
> (2*(5^3))/75
[1] 3.333333
>
```

d) 1.50

e) 1.00

f) None of the above

Question 10

Your answer is CORRECT.

Find a value of c so that P(Z ≥ c) = 0.65.

a) −0.39

b) 0.59

c) 1.39

d) −0.77

e) 0.39

f) None of the above

```
> qnorm(1-0.65)
[1] -0.3853205
>
```

Question 11

Your answer is CORRECT.

Costs for standard veterinary services at a local animal hospital follow a Normal distribution with a mean of $79 and a standard deviation of $20. What is the probability that one bill for veterinary services costs between $55 and $103?

a) 0.3849

b) 0.5000

c) 0.6151

d) 0.7699

e) 0.2301

f) None of the above

```
> pnorm(103, 79, 20) - pnorm(55, 79, 20)
[1] 0.7698607
>
```

Question 12

Your answer is CORRECT.

Current research indicates that the distribution of the life expectancies of a certain protozoan is normal with a mean of 49 days and a standard deviation of 10.2 days. Find the probability that a simple random sample of 64 protozoa will have a mean life expectancy of 54 or more days.

a) 0.2000

b) 0.0000

c) 0.0001

d) 0.6880

e) 1.0000

f) None of the above

```
> 1 - pnorm(54, 49, (10.2/sqrt(64)))
[1] 4.398719e-05
>
```

Question 13

Your answer is CORRECT.

Let x represent the hemoglobin count (HC) in grams per 100 milliliters of whole blood. The distribution for HC is approximately normal with μ = 14 for healthy adult women. Suppose that a female patient has taken 12 laboratory blood samples in the last year. The HC data sent to her doctor is listed below. We would like to know if the data indicates this patient has significantly high HC compared to the population.

$$22,\ 19,\ 14,\ 19,\ 15,\ 16,\ 21,\ 22,\ 21,\ 14,\ 20,\ 20$$

Give the p-value and interpret the results.

sigma not provided; one sample t-test

$H_0: \mu = 14$

$H_a: \mu > 14$

test statistic (from formula sheet):

$$t = \frac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}};\ \text{df} = \text{n}-1$$

to find p-value

$P(t >$ test statistic$)$

p-value -->

```
> samples = c(22,19,14,19,15,16,21,22,21,14,20,20)
> xBar = mean(samples)
> s = sd(samples)
> t = (xBar - 14)/(s/sqrt(12))
> 1 - pt(t, 11)
[1] 0.0001380108
> 0.0001380108 < 0.05
[1] TRUE
```

p-value is less than 5% significance level, so we reject the null hypothesis.

a) p = .0562; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a high HC level.

b) p = 0.0001; Based on 5% significance level, I will reject the null hypothesis and conclude this patient has a high HC level.

c) p = 0.0001; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a high HC level.

d) p = .1053; Based on 5% significance level, I will fail to reject the null hypothesis and conclude this patient does not have a high HC level.

e) p = 0.0003; Based on 5% significance level, I will reject the null hypothesis and conclude this patient has a high HC level.

Question 14

Your answer is CORRECT.

An important problem in industry is shipment damage. A pottery producing company ships its product by truck and determines that it can meet its profit expectations if, on average, the number of damaged items per truckload is fewer than 11. A random sample of 19 departing truckloads is selected at the delivery point and the average number of damaged items per truckload is calculated to be 9.4 with a calculated sample of variance of 0.49. Select a 95% confidence interval for the true mean of damaged items.

a) [10.66, 11.34]

b) [48.91, -33.04]

c) [9.063, 9.737]

d) [-0.3372, 0.3372]

e) [8.918, 9.882]

f) None of the above

we don't have population standard deviation so we use t critical value for confidence interval from the formula sheet:

One-sample t-test: $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$

```
> # sample sd is square root of sample variance
> s = sqrt(0.49)
> 9.4 + c(-1,1) * qt(1.95/2, 18) * (s/sqrt(19))
[1] 9.062611 9.737389
>
```

Question 15

Your answer is CORRECT.

An experimenter flips a coin 100 times and gets 44 heads. Test the claim that the coin is fair against the two-sided claim that it is not fair at the level α=.01.

this is a hypothesis test for proportions.

$H_o: \mu = 0.5$

$H_a: \mu \neq 0.5$

test statistic (from formula sheet):

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

```
> n = 100
> pHat = 44/n
> p = 0.50
> z = (pHat - p)/sqrt((p*(1-p))/n)
> z
[1] -1.2
>
> # p value (two tailed)
> 2 * pnorm(z)
[1] 0.2301393
> 0.2301393 < 0.01
[1] FALSE
> # p value is bigger than significance level so
> # we fail to reject the null hypothesis
>
```

a) H_0: $p = .5$, H_a: $p \neq .5$; $z = -1.20$; Fail to reject H_0 at the 1% significance level.

b) H_0: $p = .5$, H_a: $p \neq .5$; $z = -1.20$; Reject H_0 at the 1% significance level.

c) H_0: $p = .5$, H_a: $p < .5$; $z = -1.21$; Fail to reject H_0 at the 1% significance level.

d) H_0: $p = .5$, H_a: $p < .5$; $z = -1.20$; Reject H_0 at the 1% significance level.

e) H_0: $p = .5$, H_a: $p \neq .5$; $z = -1.21$; Fail to reject H_0 at the 1% significance level.

Question 16

Your answer is CORRECT.

Data for gas mileage (in mpg) for different vehicles was entered into a software package and part of the ANOVA table is shown below:

Source	DF	SS	MS	
Vehicle	4	540	270.00	<--- MSR
Error	16	204	12.75	<--- MSE
Total	20	744		

Determine the value of the test statistic F to complete the table.

a) 21.176

b) 1.378

c) 2.756

d) 0.047

e) 58.353

f) None of the above

$$F = \frac{MSR}{MSE}$$

```
> 270/12.75
[1] 21.17647
>
```

Question 17

Your answer is CORRECT.

We want to know if there is a difference between the mean list price of a three bedroom home, μ_3, and the mean list price of a four bedroom home, μ_4, $H_0 : \mu_3 = \mu_4$ versus $H_a : \mu_3 \neq \mu_4$. Suppose we get a p-value of 0.1325, which is the proper conclusion of this test?

if the p-value is less than the significance level (in this case 5%) we reject the Null Hypothesis, otherwise we fail to reject it.

```
> 0.1325 < 0.05
[1] FALSE
>
```

(its not e because if we fail to reject there is no evidence of significant diff. of mean prices of the homes)

a) We accept the null hypothesis at 5% significance. There is extremely strong evidence of a significance difference between the mean list price of three bedroom homes and the mean list price of four bedroom homes.

b) We reject the null hypothesis at 5% significance. There is no evidence of a significance difference between the mean list price of three bedroom homes and the mean list price of four bedroom homes.

c) We fail to reject the null hypothesis at 5% significance. There is no evidence of a significance difference between the mean list price of three bedroom homes and the mean list price of four bedroom homes.

d) We reject the null hypothesis at 5% significance. There is extremely strong evidence of a significance difference between the mean list price of three bedroom homes and the mean list price of four bedroom homes.

e) We fail reject the null hypothesis at 5% significance. There is extremely strong evidence of a significance difference between the mean list price of three bedroom homes and the mean list price of four bedroom homes.

f) None of the above

Question 18

Your answer is CORRECT.

A national computer retailer believes that the average sales are greater for salespersons with a college degree. A random sample of 35 salespersons with a degree had an average weekly sale of $3411 last year, while 36 salespersons without a college degree averaged $3133 in weekly sales. The standard deviations were $468 and $642 respectively. Is there evidence at the 5% level to support the retailer's belief? Select the [p-value, Decision to Reject (RH_0) or Failure to Reject (FRH_0)].

two sample t-test for means

from formula sheet:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}; \text{ df}=v$$

$H_0 : \mu_1 = \mu_2$

$H_a : \mu_1 > \mu_2$

```
> n1 = 35
> xBar1 = 3411
> s1 = 468
>
> n2 = 36
> xBar2 = 3133
> s2 = 642
>
> # test statistic
> t = (xBar1 - xBar2)/sqrt((s1^2/n1) + (s2^2/n2))
> t
[1] 2.089173
>
> # p value (right tailed)
> p = 1 - pt(t, 34)
> p
[1] 0.02212435
> # the degree of freedom is whichever n is the smallest minus 1
>
> p < 0.05
[1] TRUE
> # since p value is less than significance
> # we reject the null hypothesis
>
```

a) [p-value = 0.011, RH_0]

b) [p-value = 0.015, RH_0]

c) [p-value = 0.015, FRH_0]

d) [p-value = 0.022, FRH_0]

e) [p-value = 0.022, RH_0]

Question 19

Your answer is CORRECT.

The decline of salmon fisheries along the Columbia River in Oregon has caused great concern among commercial and recreational fishermen. The paper 'Feeding of Predaceous Fishes on Out-Migrating Juvenile Salmonids in John Day Reservoir, Columbia River' (Trans. Amer. Fisheries Soc. (1991: 405-420) gave the accompanying data on y = maximum size of salmonids consumed by a northern squaw fish (the most abundant salmonid predator) and x = squawfish length, both in mm. Here is the computer software printout of the summary:

Coefficients:				
	Estimate	Std. Error	t value	Pr(> |t|)
(Intercept)	−92.040	16.706	−5.509	0.000
Length	0.710	0.041	17.275	0.000

x ---> Length

Using this information, give the equation of the least squares regression line and the 95% confidence interval for the slope.

a) $\hat{y} = 16.706x - 92.040$; [0.642, 0.778]

b) $\hat{y} = -92.040x + 0.710$; [−125.986, −58.094]

c) $\hat{y} = 0.710x + 16.706$; [−124.783, −59.297]

d) $\hat{y} = 16.706x + 0.041$; [0.371, −57.839]

e) $\hat{y} = 0.710x - 92.040$; [0.626, 0.794]

f) None of the above

Question 20

Your answer is CORRECT.

Which of the following would be the LSRL for the given data?

x	1	8	8	11	16	17
y	21	28	29	41	32	43

a) $\hat{y} = 1.136\,x + 20.78$

b) $\hat{y} = -1.136\,x + 20.78$

c) $\hat{y} = -20.78\,x + 1.136$

d) $\hat{y} = 20.78\,x + 1.136$

e) None of the above

look here -->

```
> x = c(1, 8, 8, 11, 16, 17)
> y = c(21, 28, 29, 41, 32, 43)
> y.lm = lm(y~x)
> summary(y.lm)

Call:
lm(formula = y ~ x)

Residuals:
      1       2       3       4       5       6
-0.9171 -1.8713 -0.8713  7.7197 -6.9619  2.9018

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   20.781      4.800   4.330   0.0124 *
x              1.136      0.417   2.725   0.0527 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.513 on 4 degrees of freedom
Multiple R-squared:  0.6499,	Adjusted R-squared:  0.5624
F-statistic: 7.427 on 1 and 4 DF,  p-value: 0.0527

>
```