1. A study is conducted on students taking a statistics class. Several variables are recorded in the survey. Select **all** variables that are quantitative (numerical).

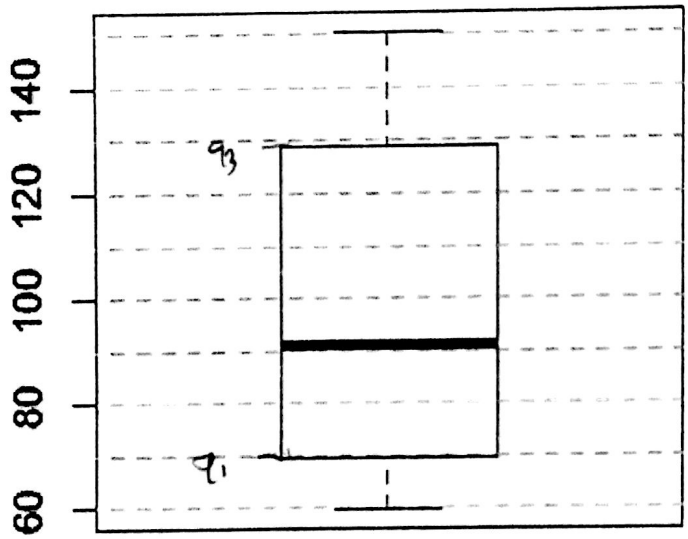
A) Type of car the student owns.

B) Number of credit hours taken during that semester.

C) The time the student waited in line at the bookstore to pay for his/her textbooks.

D) Home state of the student.

Problems 2 – 4. The *boxplot* given below is based on **60** measurements.



2. About 25% of the measurements are **less** than

(A) 70 (B) 90 (C) 110 (D) 130 (E) 150

3. About 30 of the measurements are **more** than

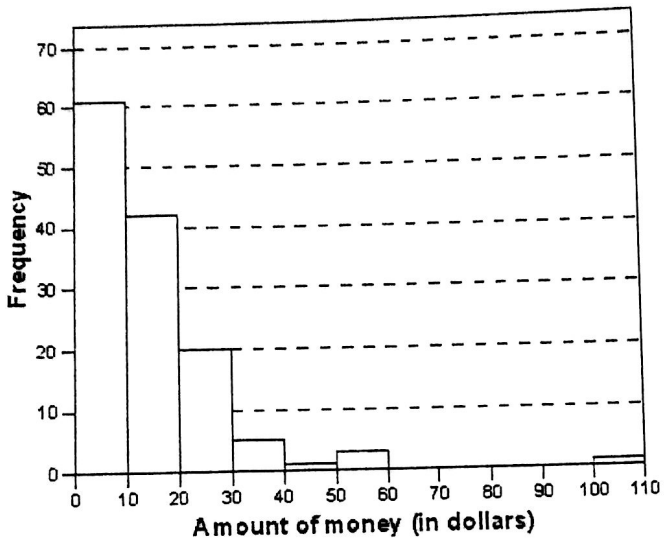
(A) 70 (B) 90 (C) 110 (D) 130 (E) 150

4. What is the IQR for the given data?

(A) 50 (B) 60 (C) 70 (D) 80 (E) 90

$IQR = q_3 - q_1 = 130 - 70 = 60$

Problem 5. In a statistics class with 136 students, the professor records how much money each student has in their possession during the first class of the semester. The histogram shown below represents the data he collected:



5. Which of the following description(s) is/are correct regarding the shape of the histogram? Choose **all** relevant answers.

A) Skewed right
B) Skewed left
C) Symmetric
D) An outlier is present.
E) Unimodal
F) Bimodal

Problems 6 – 7. The asking prices (in thousands of dollars) for a sample of 13 houses currently on the market in Neighborville are listed below. For convenience, the data have been ordered.

175 199 205 234 259 275 299 304 317 345 355 384 549

min, Q_1, $\tilde{x}$, Q_3, max
175, 234, 299, 345, 549

6. What is the five-number summary?

A)	175	234	299	345	549
B)	175	219.5	299	350	549
C)	175	219.5	299	350	384
D)	175	234	299	331	549

7. Use the 1.5 × *IQR* rule to determine if there are any outliers present. What is/are the value(s) of the outlier(s)?

A) No outliers present
B) One outlier: 175
C) One outlier: 549
D) Two outliers: 175 and 549

$IQR = Q_3 - Q_1 = 345 - 234 = 111$

$Q_1 - 1.5IQR < 234 - 1.5(111)$
$= 67.5$

$Q_3 + 1.5IQR = 345 + 1.5(111)$
$= 511.5$

$549 > 511.5$

Problems 8 – 12. The distribution of the weights of certain items can be approximated by a Normal distribution with mean 90 grams and a standard deviation of 2 grams.

Using the Empirical Rule [i.e. 68-95-99.7 Rule],

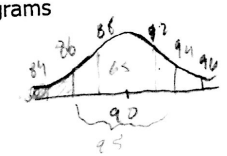
$\sigma = 2$ $\mu = 90$

8. What percentage of the items will weigh less than 90 grams
A) 68%
B) 50%
C) 32%
D) 16%

9. What percentage of the items will weigh between 84 grams and 86 grams
A) 13.5%
B) 4.7%
C) 2.5%
D) 2.35%

$2.5 - .15$

2.35



Using the table provided,

10. What proportion of the items will weigh more than 86.25g
A) 0.0301
B) 0.0375
C) 0.9699
D) 0.9750

$z = \frac{86.25 - 90}{2} = -1.875$

$1 - .0304 = .9696$

11. What proportion of the item weigh between 91.28g and 92.06g
A) 0.8485
B) 0.1096
C) 0.0548
D) 0.1161

$z_1 = \frac{91.28 - 90}{2} = .64$

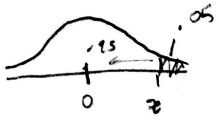
$z_2 = \frac{92.06 - 90}{2} = 1.03$

$.8485 - .7389 = .1096$

12. How much must an item weigh to be amongst the largest 5% of all the items?
A) 93.29g
B) 86.71g
C) 93.92g
D) 86.08g

$1.645 = \frac{x - 90}{2}$

$x = 93.29$



$z = 1.645$

13. A continuous variable x has the density function,

$$f(x) = c(1 - x^2) \qquad 0 \le x \le 1$$

Determine the value of c.

A) 0.67 B) 1.5 C) 2 D) 2.5

$\int_0^1 c(1 - x^2)\,dx = 1$

$c\int_0^1 (1 - x^2)\,dx = 1$, $c\left[x - \frac{1}{3}x^3\right]_0^1 = 1$ $c\left(1 - \frac{1}{3}\right) = 1$

$c(2/3) = 1$ $c = 3/2$

– 0

Problems 14 - 15. A continuous variable x is said to have a *uniform* distribution if the density function is given by,

$f(x) = \frac{1}{b-a}$ for $a < x < b$ and 0 otherwise.

Suppose the time (min) taken by a clerk to process a certain application form has a uniform distribution with a = 4 and b = 6.

14. In the long run, what proportion of forms will take between 4.5 min and 5.5 min to process?

$\frac{1}{6-4} = \frac{1}{2}$ $\frac{1}{5.5-4.5} = 1$

A) 0.25 B) 0.5 C) 0.75 D) 1

15. What value separates the best (i.e. fastest) 10% of all processing times from the remaining 90%?

$.1(2) = .2$

A) 4.2 B) 4.4 C) 5.6 D) 5.8

5.8

16. Identify a legitimate mass function for a discrete variable y,

A) p(0) = .2, p(1) = .4, p(2) = .1

B) p(y) = 0.45(2-y) for y = 0, 1, 2 $.9 + .45 \times$

C) p(0) = p(1) = 0.2, p(3) =0.1, p(4) = p(5) = 0.25

D) p(y) = 0.2(3-y) for y = 0, 1, 2, 3, 4. $.6 + .4 + .2 + .2$

Problems 17 – 19. A college professor always finishes his lectures within 2 minutes after the bell rings to end the period and the end of the lecture. Let X = the time that elapses between the bell and the end of the lecture and suppose the pdf of X is

$$f(x) = \begin{cases} kx^2 & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$\int_0^2 kx^2 dx = 1 = \left(\frac{k}{3}x^3\right)_0^2 = 1$

$\frac{k}{3}(8) = 1 \quad k = 3/8$

17. Find the mean, μ, of this distribution.

$\mu = \int_0^2 x \, kx^2$

$\mu = k\left[\frac{1}{4}x^4\right]_0^2 = k[4] = 4(3/8)$

$= 1.5$

A) 1.5
B) 0.125
C) 0.375
D) any positive value greater than 2

18. What is the probability that the lecture ends within 1 minute of the bell ringing?

$\int_0^1 \frac{3}{8}x^2 dx$

$\left[\frac{1}{8}x^3\right]_0^1 = \frac{1}{8}$

A) 1.5
B) 0.125
C) 0.375
D) any positive value greater than 2

19. What percentage of the time does the lecture continue for at least 90 seconds beyond the bell?

A) 57.81%
B) 92.67%
C) 23.99%
D) 49.32%
E) None of the above

Problems 20 – 21. A student must pass through 3 traffic lights on her way to class. A mass function for the number of red lights (x) she hits on her way is given below.

x - # of red lights	0	1	2	3
p(x)	0.4	0.3	0.2	0.1

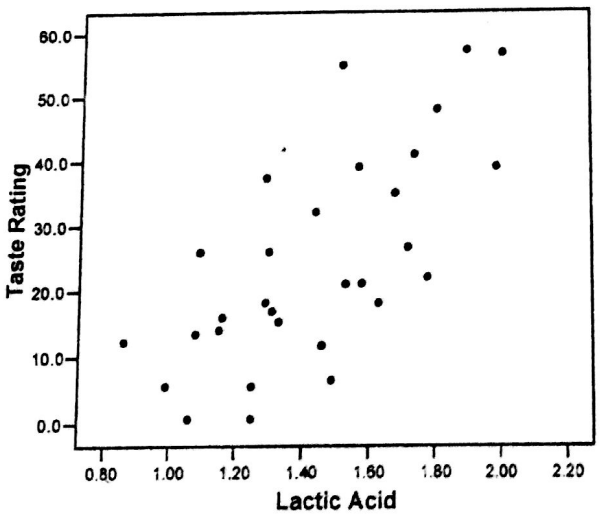
20. Find the mean of x.

(A) 0.8 (B) 1.0 (C) 1.2 (D) 1.4

21. Find the standard deviation of x.

(A) 0.8 (B) 1.0 (C) 1.2 (D) 1.4

22. The taste of matured cheese is related to the concentration of several chemicals in the final product. A scatterplot of the observed data is shown below:



What is a plausible value for the correlation between lactic acid concentration and taste rating?

A) 0.7
B) 0
C) 0.07
D) –0.7

23. True or False. A correlation of –0.9 indicates a stronger linear relationship than a correlation of 0.8.

(A) TRUE (B) FALSE

Problems 24 – 28. Suppose you want to know how salaries (in thousands of dollars per year) are related to years of experience for your chosen career. You collect data for a random sample of 100 people with this type of job who have had from 0 to 20 years of experience. The relationship is linear with a correlation of 0.904 and you determine the regression equation to be:

salary = 45.59 + .798 years.

24. For this equation, the number 45.59 indicates
(A) The y-intercept is meaningless because of extrapolation.
(B) An individual with no experience in this career will earn $45.59 per hour.
(C) An individual with no experience in this career will earn $45,590 per year
(D) An individual with between 0 to 20 years of experience will earn $49,590 per year.

25. Shawn Manning is in this profession and has 15 years of experience. Use the regression equation above to predict his salary
(A) $45 590
(B) $57 560
(C) We can't use this regression equation as this is extrapolation.
(D) $79 800

26. Shawn Manning (who has 15 years working experience) informs you that his actual salary is $70 000 per year. Calculate the residual.

(A) $24 410
(B) -$24 410
(C) $12 440
(D) -$9 800

70000 − 57560
= 12440

27. True or False. There is a positive association between the years of experience and the salary for this type of job.

(A) True (B) False

$r^2 = ?$

28. What percent of the variation in the values of the salary is accounted for by the linear relationship between salary and years of experience?
A) 66.7%
B) 81.7%
C) 90.4%
D) Cannot be determined from the information given.

Problems 29 – 30. The following is partial output from MINITAB.

```
Regression Analysis: y versus x
The regression equation is
y = a + b x


Predictor       Coef    SE Coef      T      P
Constant      0.8533     0.4802   1.78  0.106
x           0.028194   0.008300   3.40  0.007


S = 0.845271   R-Sq = 53.6%   R-Sq(adj) = 48.9%
```

29. What is the value of b (the slope)?

A) 0.0083 B) 0.0282 C) 0.8533 D) 0.4802

30. What is the correlation between x and y?

A) -0.845 B) 0.536 C) 0.732 D) -0.489

TABLE A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

TABLE A Standard normal probabilities *(continued)*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Health Plan	1	2
1	27%	14%
2	24%	35%

a) Four events
① Health Plan 1 and Dental Plan 1
② Health Plan 1 and Dental Plan 2 ✓
③ Health Plan 2 and Dental Plan 1
④ Health Plan 2 and Dental Plan 2

b) P{restrictive dental plans}
P{HP1, DP1} + P{HP2, DP1}
= .27 + .24
= **.51 = 51%** ✓

c) P{more restrictive plan of each type}
P{HP1, DP1}
= **.27 = 27%** ✓

d) P{flexible dental and restrictive health}
P{HP1, DP2}
= **.14 = 14%** ✓

	Color			
Transmission Type	White	Blue	Black	Red

A - (auto. transmission) B - black C - white

a) P(A), P(B), and P(A and B)?

$P(A) = .13 + .10 + .11 + .11$

$\boxed{P(A) = .45 = 45\%}$

$P(B) = .11 + .15$

$\boxed{P(B) = .26 = 26\%}$

$\boxed{P(A \text{ and } B) = .11 = 11\%}$

b) $P(A|B)$ and context?

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

$$= \frac{.11}{.26}$$

$$\boxed{P(A|B) = 0.4231}$$

This is the probability of A given B. In this case it is the probability of the sports car being equipped with automatic transmission given that the car is black. So there is a 42.31% chance that a car is equipped with automatic transmission given that it is black.

c) $P(B|A)$ and context?

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

$$= \frac{.11}{.45}$$

$$\boxed{P(B|A) = .2444}$$

transmission, there is a 24.4% chance that a car is black given that it is equipped with automatic transmission.

d) P(A|C) and context?

$$P(A|C) = \frac{P(A \text{ and } C)}{P(C)}$$

$P(A \text{ and } C) = .13$

$P(C) = .28$

$$P(A|C) = \frac{.13}{.28}$$

$$\boxed{P(A|C) = .4643}$$

This is the probability that a chosen sports car is equipped with automatic transmission given that it is white. There is a 46.43% chance that the car is equipped with automatic transmission given

3.

x	0	1	2	3
p(x)	.8	.1	.05	.05

a) $\mu = \Sigma x \cdot p(x)$

$\mu = 0(.8) + 1(.1) + 2(.05) + 3(.05)$

$\mu = .1 + .1 + .15$

MEAN $\boxed{\mu = .35}$ ✓

STANDARD DEVIATION

$\sigma^2 = \Sigma(x-\mu)^2 \cdot p(x)$

$\sigma^2 = (0-.35)^2(.8) + (1-.35)^2(.1) + (2-.35)^2(.05) + (3-.35)^2(.05)$

$\sigma^2 = .098 + .04225 + .136125 + .351125$

$\sigma^2 = .6275$

$\boxed{\sigma = .7921}$ ✓

b) $n = 64$

$\mu_{\bar{x}} = \mu$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

mean: $\boxed{\mu_{\bar{x}} = .35}$ ✓

$\sigma_{\bar{x}} = \frac{.7921}{\sqrt{64}}$

standard deviation: $\boxed{\sigma_{\bar{x}} = .0990}$

c) P{average number of flaws per grill exceeds 1}

$P\{\bar{x} > 1\}$

$\bar{x} \sim N(.35, .0990)$

$z = \frac{1 - .35}{.0990} = 6.56565$ ⇐ ? # of off the chart.

$\boxed{P\{\bar{x} > 1\} = 0\%}$ ✓

There is approximately 0% chance that in a group of 64 the average number of flaws per grill will exceed 1.

4.

Species	n	$\bar{x}$	S
Mugil liza	56	9.15	1.27
Pogonias cromis	61	3.08	1.71

a) 95% two sided confidence interval for population mean concentration for Mugil liza

$\bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = (a,b)$ $\quad \sigma \approx S$

$\alpha = 1 - C = .05$

$z_{.025} = 1.96$

$9.15 \pm 1.96\left(\frac{1.27}{\sqrt{56}}\right) = (a,b)$

$a = 8.8174$

$b = 9.4826$

$\boxed{(a,b) = (8.8174, 9.4826)}$

b) 99% confidence interval for Pogonias cromis species

$\bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = (a,b)$ $\quad \sigma \approx S$

$\alpha = 1 - .99 = .01$

$z_{.005} = 2.575$

$3.08 \pm 2.575\left(\frac{1.71}{\sqrt{61}}\right) = (a,b)$

$a = 2.5162$

$b = 3.6438$

$\boxed{(a,b) = (2.5162, 3.6438)}$

STATISTICS FOR ENGINEERS & SCIENTISTS
ASSESSMENT

FIRST NAME LAST NAME

1. In an exercise on reducing cholesterol, a researcher sampled 50 people from a local gym who exercise regularly and 50 people from the surrounding community who do not exercise regularly. Each subject reported to a clinic to have their cholesterol measured. The subjects were unaware of the purpose of the study, and the technician measuring the cholesterol was not aware of whether the subject exercises regularly or not.

What type of study is this?

A. An observational study ✓ B. An experimental study

2. Suppose that $P(A \cap B) = 0.32$ and $P(B) = 0.56$. Find $P(A \mid B)$.

A. 0.8800 B. 0.1792 C. 0.5714 D. 0.3721

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

$= .5714$

Problems 3-6. In a CD company two machines produce CD's. Machine 1 (M1) produces only 1% of the CDs. Of the CDs produced by M1, 29% are defective whilst 4% of the CDs produced by machine 2 (M2) are defective. Let D denote defective CDs and D^c denote the non-defective CDs.

$P(M_1) = .01$ $P(M_2) = .99$ $P(D|M_1) = .29$ $P(D|M_2) = .04$

3. Find $P(D^c \mid M1)$.

A. 0.0100 B. 0.7100 ✓ C. 0.9900 D. 0.0400

4. Find $P(M1 \cap D^c)$.

A. 0.0396 B. 0.7100 C. 0.0071 ✓ D. 0.9504

5. Find $P(D^c)$.

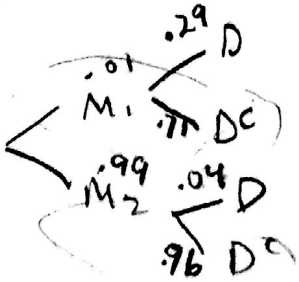
$.0071 + P(M_2 \cap D^c) = .0071 + .9504 = .9575$

A. 0.8321 B. 0.9467 C. 0.8540 D. 0.9575 ✓

6. Find $P(M1 \mid D^c)$.

A. 0.00152 B. 0.007415 ✓ C. 0.003896 D. 0.001761

$P(M_1|D^c) = \frac{P(M_1 \cap D^c)}{P(D^c)} = \frac{.0071}{.9575}$



Problems 7-8. Suppose that the events A and B are **independent**. Furthermore, suppose that P(A) = 0.37 and P(B) = 0.45.

7. Find P (A ∩ B).

   A. 0.8200 B. 0.8222 C. 0.1665 D. 0.6535

8. Find P (A ∪ B). $= P(A) + P(B) = .82 - .1665$

   A. 0.8200 B. 0.8222 C. 0.1665 D. 0.6535

9. Let A and B be two events where P(A) = 0.55 and P(B) = 0.63. TRUE or FALSE, it is possible that events A and B are **disjoint**. $P(A) + P(B) = 1.18$

   A. TRUE B. FALSE

Problems 10 -11. A specific device's lifetime (hrs) has a distribution with mean $\mu = 50$ and variance $\sigma^2 = 225$. From this population a simple random sample of n observations is to be selected and the mean of the sample values calculated.

If the population variable is known to be Normally distributed and the sample size used is to be n = 16,

10. What is the probability that the mean lifetime of 16 randomly selected devices will be between 48.35 and 55.74?

    $(48.35, 55.74) = 50 \pm z_{\alpha/2}\left(\frac{15}{4}\right)$

    $z_{\alpha/2} = 1.5333$

    A. 0.393
    B. 0.607
    C. 0.937
    D. 0.330
    E. Not within ± 0.010 of any of the above.

11. What is the probability that a randomly selected device's lifetime is between 48.35 and 55.74?

    A. 0.4562
    B. 0.6480
    C. 0.1918
    D. 0.3300
    E. Not within ± 0.010 of any of the above.

Problems 12-13. In a certain population, 42% have type O blood.

12. What is the probability that in a random selection of 4 individuals no one has type O blood?

    $(.58)^4$

    A. 0.1132 B. 0.1575 C. 0.1385 D. 0.0983 E. 0.1785

13. What is the probability that in a random selection of 5 individuals at least one has type O blood?

    A. 0.9008 B. 0.9344 C. 0.8840 D. 0.9449 E. 0.9155

$1 - p(\text{zero})$

$1 - (.58)^5$

Problems 14 – 15. Assume that sample data, based on two independent samples of size 25, give us $\bar{x}_1 = 505$, $\bar{x}_2 = 515$, $s_1 = 23$, and $s_2 = 28$.

14. What is a 95% confidence interval for $\mu_2 - \mu_1$?

A. (-2.40, 22.40)
B. (-4.57, 24.57)
C. (-4.96, 24.96)
D. (5.79, 14.21)
E. None of the above

$n = 25$ $df = 24$

$(a,b) = (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{23^2}{25} + \frac{28^2}{25}}$

$t_{\alpha/2} = 2.064$ 2.797

$(a,b) = (505 - 515) \pm 2.064\sqrt{\frac{23^2}{25} + \frac{28^2}{25}}$

-10

$(a,b) = (-4.96, 24.96)$

15. Determine which of the following statements is TRUE.

A. Based on the confidence interval, we can conclude, at the 5% significance level, that there is no difference between the two population means, μ_2 and μ_1.
B. The margin of error for the difference between the two sample means would be larger if we were to take larger samples.
C. If a 99% confidence interval were calculated instead of the 95% interval, it would include fewer values for the difference between the two population means.
D. None of the above

Problems 16 – 17. You wish to test if there are any differences in the prices of apartments in two neighboring towns. You take a simple random sample of 12 apartments in town A and calculate the average price of these apartments. You repeat this for 15 apartments in town B. Let μ_1 represent the true average price of apartments in town A and μ_2 the average price in town B.

$n_A = 12$
$n_B = 15$

16. What would be the hypotheses for this problem?

A. H_0: $\mu_1 = \mu_2$ versus H_a: $\mu_1 < \mu_2$
B. H_0: $\mu_1 = \mu_2$ versus H_a: $\mu_1 > \mu_2$
C. H_0: $\mu_1 = \mu_2$ versus H_a: $\mu_1 \neq \mu_2$

17. Suppose we were to use the *t* test with the conservative estimate for the degrees of freedom. The *t* statistic for comparing the mean prices is 2.1. What can we say about the value of the *P*-value?

$t = 2.1$ $df = 11$

A. P-value < 0.01
B. 0.025 < P-value < 0.05 $\times 2$ $-1/2$
C. 0.05 < P-value < 0.10
D. P-value > 0.10

18. Complete the following sentence. A type II error occurs when

A. the null hypothesis is not rejected and the null hypothesis is true.
B. the null hypothesis is rejected and the null hypothesis is false
C. the null hypothesis is not rejected and the null hypothesis is false.
D. the null hypothesis is rejected and the null hypothesis is true.

19. Suppose that you are performing a t-test. For H_a: $\mu \neq 1$ and T = 2.5, the p-value is the area under the density function
A. to the left of 2.5.
B. to the right of 2.5.
C. to the left of -2.5 and the right of 2.5.

20. The heights of young American women, in inches, are Normally distributed with mean μ and standard deviation $\sigma = 2.4$. The margin of error for a 99% confidence interval is desired to be within 1 inch. What should be the sample size? C=.99

A. 7 B. 22 C. 38 D. 39 $\sigma = 2.4$ $n = \left(\frac{z_{\alpha/2}\,\sigma}{B}\right)^2 = 38.1924 \Rightarrow 39$

Problems 21 – 22. A research paper reports that the p-value for the one-sample t-test of H_0: $\mu = 1$ versus H_a: $\mu \neq 1$ is 0.06. The test statistic is -2.3. [Hint: It might help to sketch the density curve of the test statistic].

21. The p-value for the test H_0: $\mu = 1$ versus H_a: $\mu > 1$ is:
A. 0.03 B. 0.06 C. 0.94 D. 0.97

pvalue = .06

22. The p-value for the test H_0: $\mu = 1$ versus H_a: $\mu < 1$ is:
A. 0.03 B. 0.06 C. 0.94 D. 0.97

23. The square footage of the several thousand apartments in a new development is advertised to be 1250 square feet, on average. A tenant group thinks that the apartments are smaller than advertised. Let μ represent the true average area (in square feet) of these apartments. What are the appropriate null and alternative hypotheses?

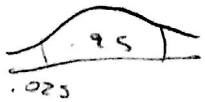
A. H_0: $\mu = 1250$ vs. H_a: $\mu < 1250$

B. H_0: $\mu = 1250$ vs. H_a: $\mu \neq 1250$

C. H_0: $\mu = 1250$ vs. H_a: $\mu > 1250$

Problems 24 – 27. A national survey carried out in 2000 showed that 10% of student loan borrowers had loans of more than $30,000. A more recent survey of 1280 student loan borrowers found that 192 had loans totaling more than $30,000 for their undergraduate education. A researcher who conducted this recent survey suspects that the proportion of borrower with loans of over $30,000 has gone up in the last 10 years.

24. Find the margin of error for a two-sided 95% confidence interval for π.



A. 0.01956 B. 0.01643 C. 0.15 D. 0.01754 E. None of these

192/1280
$\pi = .15$

1.95

7. $z = 2.81$ $z\sqrt{\frac{\pi(1-\pi)}{n}} = .028$?

$\frac{.028}{2} = .014$

25. State the hypotheses to test the researcher's suspicion.

A. H_0: $\pi = 0.10$ vs. H_a: $\pi \neq 0.10$

B. H_0: $\pi = 0.10$ vs. H_a: $\pi > 0.10$

C. H_0: $\pi = 0.10$ vs. H_a: $\pi < 0.10$

26. Suppose Minitab software gives a p-value of 0.0157 for this test. Is there evidence to support the researcher's suspicion at 5% significance level?

A. Yes B. No

p value $< \alpha$
reject H_0

27. The researcher wishes to conduct a different survey. What sample size would they need for a 90% margin of error that does not exceed 1%.

A. 68 B. 6765 C.6766 D. 6764

28. A SRS of 100 postal employees is used to test if the average time postal employees have worked for the postal service has decreased from the value of 7.5 years recorded 20 years ago. The sample mean was 7 years with a standard deviation of 2 years. Assume the distribution of the time the employees have worked for the postal service is approx. Normal. At 1% significance level, what decision is appropriate?

$\bar{x} = 7$ $s = 2$ $n = 100$
$\alpha = .01$
$z = \frac{.5}{2/10} = 2.5$
.006 z

A. Reject H_0 B. Fail to reject H_0

29. Suppose the 98% confidence interval for μ_1-μ_2 is (0.0345, 3.986). Without recalculating anything else and based solely on this confidence interval, would you reject H_0: $\mu_1 = \mu_2$ in favor of H_a: $\mu_1 \neq \mu_2$ at 5% significance level?

A. Yes B. No C. Cannot tell

30. The time needed for college students to complete a certain paper-and-pencil maze follows a Normal distribution with a mean of 30 seconds and a standard deviation of 3 seconds. You wish to see if the mean time μ is changed by vigorous exercise, so you have a group of nine college students exercise vigorously for 30 minutes and then complete the maze. Assume that σ remains unchanged at 3 seconds. The hypotheses you decide to test are H_0: $\mu = 30$ versus H_a: $\mu \neq 30$. Suppose it takes the nine students an average of = 32.05 seconds to complete the maze. At the 2% significance level, what can you conclude?

$\bar{x} = 30$
$n = 9$
32.05 - 30

A. H_0 should be rejected because the P-value is less than 0.02.
B. H_0 should not be rejected because the P-value is greater than 0.02.
C. Ha should be rejected because the P-value is less than 0.02.
D. Ha should not be rejected because the P-value is greater than 0.02.

STAT3010: STATISTICS FOR ENGINEERS AND SCIENTISTS
PRACTICE QUESTIONS

1. Which of the following are examples of a variable?
   A) Gender of a high school graduate CAT.
   B) Number of major credit cards a person has JUAN.
   C) Type of automobile transmission CAT.
   D) All of the above

2. A stem-and-leaf display describes two-digit integers between 30 and 85. For one of the classes displayed, the row appears as 6|357. What numerical values are being described?
   A) 36, 56, and 76
   B) 50, 60, and 70
   C) 63, 65, and 67
   D) 35 and 57

3. Which of the following statements regarding histograms are correct?
   A) A unimodal histogram is one that rises to a single peak and then declines, whereas a bimodal histogram is one that has two different peaks.
   B) A unimodal histogram is positively skewed if the right or upper tail is stretched out compared to the left or lower tail.
   C) A unimodal histogram is negatively skewed if the left or lower tail is stretched out compared to the right or upper tail.
   D) A histogram is symmetric if the left half is a mirror image of the right half.
   E) All of the above

4. A histogram that is negatively skewed has
   A) longer tail extends to the right
   B) longer tail extends to the left
   C) shorter tail extends to the right
   D) shorter tail extends to the right
   E) left half identical to right half

5. A histogram that is positively skewed has
   A) longer tail extends to the right
   B) longer tail extends to the left
   C) shorter tail extends to the right
   D) shorter tail extends to the right
   E) left half identical to right half

6. The probability mass function of a discrete random variable x is defined as $p(x) = x/10$ for $x = 0, 1, 2, 3$, and 4. Then, the probability that x is at most 3 is
   A) .10
   B) .30
   C) .60
   D) .90

   $x \leq 3$
   $1/10 + 2/10 + 3/10 + 4/10$
   $.1 + .2 + .3$

7. For a discrete random variable x, if $P(x \text{ is at most } 0) = .3$, $P(x \text{ is at most } 1) = .7$, $P(x \text{ is at most } 2) = .9$, and $P(x \text{ is at most } 3) = 1.0$, then the value of the probability mass function $p(x)$ at $x = 1$ is
   A) .30
   B) .40
   C) .20
   D) .80

   $.7 - .3$

8. If the probability density function of a continuous random variable x is

$$f(x) = \begin{cases} .5x & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

then $P(1 \le x \le 1.5)$ is

A) .5625
B) .3125
C) .1250
D) .4375

9. Which of the following standard normal probabilities are not correct?

A) $P(z \le -1.25) = .8944$
B) $P(z > 1.25) = .1056$
C) $P(z \le 1.25) = .8944$
D) $P(-1.25 \le z \le 1.25) = .7888$

10. If x is a normally distributed random variable with a mean of 10 and standard deviation of 4, then the probability that x is between 6 and 16 is

A) .9332
B) .7745
C) .6587
D) .0668
E) .8413

11. If x is a normally distributed random variable with a mean of 25 and a standard deviation of 8, then the probability that x exceeds 20 is approximately

A) .7357
B) .2643
C) .6250
D) .3750
E) None of the above answers are correct

12. Which of the following statements are true regarding the areas beneath any normal curve?

A) About 68.3% of the area is in the interval $\mu - \sigma$ to $\mu + \sigma$.
B) About 95.5% of the area is in the interval $\mu - 2\sigma$ to $\mu + 2\sigma$.
C) Nearly all of the area (about 99.7%) is in the interval $\mu - 3\sigma$ to $\mu + 3\sigma$.
D) All of the above are true statements.
E) None of the above is true statement.

13. Let z_1 be a z score that is unknown but identifiable by position and area. If the area to the right of z_1 is 0.8413, then the value of z_1 must be

A) 1.00
B) -1.00
C) 0.00
D) 0.41
E) -0.41

14. The measure most affected by outliers is the:
A) mean
B) median
C) trimmed mean
D) All of the above

15. The average score for a class of 25 students was 75. If the 15 female students in the class averaged 70, then the male students in the class averaged:
A) 85.0
B) 82.5
C) 77.5
D) 75.0
E) 70.0

16. Which of the following statements are true for the following data values: 9, 7, 8, 6, 9, 10, and 14?
A) The mean and median are equal
B) The mean is larger than the median
C) The mean is smaller than the median
D) The mean and the 10% trimmed mean are equal
E) None of the above statements is true

17. Which of the following statements are not true if the lower and upper quartiles of a data set are 26 and 34, respectively?
A) The interquartile range is 8
B) The interquartile range is 30
C) Any observation farther than 12 from the closet quartile is an outlier
D) Any outlier that is less than 24 from the closet quartile is considered mild outlier
E) All of the above

18. Which of the following summary measures is affected most by outliers?
A) The lower quartile
B) The median
C) The upper quartile
D) The interquartile range
E) None of the above

19. Which of the following statements are not true?
A) The objective of scatter plots is to get an idea about the relationship between two variables so that we can gain information about one of them through knowing values of the other(s).
B) Saying that variables x and y are positively linearly related means that y increases (decreases) as x increases (decreases).
C) Saying that variables x and y are negatively linearly related means that y increases (decreases) as x decreases (increases).
D) All of the above statements are true.
E) None of the above statements are true.

20. Which of the following statements are not true?

A) In studying the relationship between two variables x and y, the independent variable is represented by x.
B) In studying the relationship between two variables x and y, the dependent variable is represented by y.
C) A first step in studying the relationship between two variables x and y is to construct a scatter plot for the (x, y) values.

D) None of the above statements is true
E) All of the above statements are true.

21. Which of the following statements are not true about the sample correlation coefficient r?
A) The value of r depends on which of the two variables under study is labeled x and which is labeled y.
B) The value of r is independent of the units in which x and y are measure.
C) The value of r is always between -1 and +1, inclusive.
D) The value of $r = 1$ if all (x_i, y_i) pairs lie on a straight line with positive slope.
E) The value of $r = -1$ if all (x_i, y_i) pairs lie on a straight line with negative slope.

22. A data set consists of 15 pairs of observations $(x_1, y_1), (x_2, y_2), \ldots\ldots (x_{15}, y_{15})$. If each x_i is replaced by $3x_i$ and if each y_1 is replaced by $4y_i$, then the sample correlation coefficient r
A) increases by 3/15
B) increases by 4/15
C) remains unchanged
D) decreases by 3/15
E) decreases by 4/15

23. A data set consists of 20 pairs of observations $(x_1, y_1), (x_2, y_2), \ldots\ldots (x_{20}, y_{20})$. If each x_i is replaced by $x_i - 1$ and if each y_i is replaced by $y_i - 2$, then the sample correlation coefficient r
A) decreases by .05
B) decreases by .10
C) increases by.05
D) increases by .10
E) remains unchanged

24. Which of the following statements are not true?
A) The correlation coefficient r is a measure of how strongly related x and y are in the observed sample, while the correlation coefficient ρ is a measure of how strongly related x and y are in the population
B) The population correlation coefficient ρ is a number between 0 and 1
C) The population correlation coefficient $\rho = +1$ or -1 if and only if all (x, y) pairs in the population lie exactly on a straight line, so ρ measures the extent to which there is a linear relationship in the population.
D) All of the above statements are true
E) None of the above statements are true.

25. The following values are listed as coefficients of correlation (r). Which one indicates a positive relationship between the variables x and y?
A) 0.0
B) -0.7
C) 0.8
D) 1.4
E) -1.3

26. If all the points in a scatter diagram lie on a straight line, then the coefficient of correlation r must be:

A) +1.0
B) –1.0
C) either +1.0 or –1.0
D) 0
E) any positive value