STAT 328
Summer 2004
Lab 1
Solutions

1.20

(a)

```
1 | 0 4 4 4 5
2 | 0 2 2 6 6 7
3 | 4 6
4 | 2 8
```

```
1 | 0 4 4 4
1 | 5
2 | 0 2 2
2 | 6 6 7
3 | 4
3 | 6
4 | 2
4 | 8
```

The second plot, with the split stems is preferred because it makes the skewness more obvious.

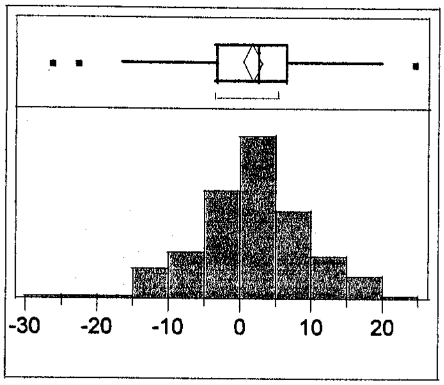
(b) right-skewed

(c) Center: Median = 22
Spread: 10 – 48

I would comment about the typical decline being around 22%, but could be as high as 48% or as low as 10%

1.53

Monthly Percent Returns



Stem and Leaf

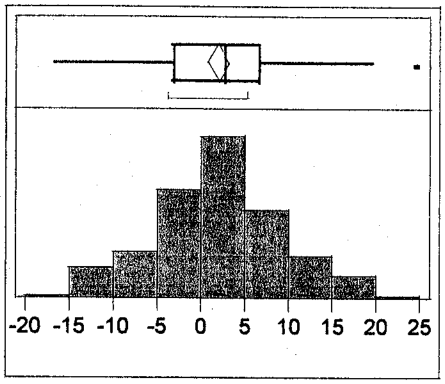
Stem	Leaf	Count
2		
2	4	1
1	778999	6
1	0000111122234	13
0	555555566667777788889999	24
0	1111111112222233333333334444444444444	37
-0	44444333333332222221000000	26
-0	998887776655555	15
-1	42110000	8
-1	65	2
-2	3	1
-2	7	1

-2|7 represents -27

(a) Above is a histogram and a stem-and-leaf plot. Notice on the stem-and-leaf plot that the data were rounded to the nearest integer value. The histogram is preferred for a large data set such as this one.

(b) The values of the two outliers are are –26.6 and –22.9.

Distributions
Monthly Percent Returns



Quantiles

100.0%	maximum	24.40
99.5%		24.40
97.5%		19.04
90.0%		11.18
75.0%	quartile	6.63
50.0%	median	2.65
25.0%	quartile	-3.20
10.0%		-8.37
2.5%		-13.61
0.5%		-16.50
0.0%	minimum	-16.50

Moments

Mean	1.95
Std Dev	7.6022847
Std Err Mean	0.661694
upper 95% Mean	3.2589885
lower 95% Mean	0.6410115
N	132

The median is at 2.65 with a spread of –16.5 to 24.4. The shape is roughly symmetric.

(c)
Moments

Mean	1.5514925
Std Dev	8.2181052
Std Err Mean	0.7099362
upper 95% Mean	2.9557188
lower 95% Mean	0.1472663
N	134

You can see from the above JMP output that the mean return with the outliers included is about 1.55% with a standard deviation of 8.21. If you were to invest $100 dollars at the beginning of the month and got the mean return, you would have $101.55 at the end of the month.

(d) If you were to invest $100 dollars and receive the worst realized return of –26.6%, you would be left with $73.40 at the end of the month. Looking at the JMP output in part (b) we see a mean return of 1.95% and a standard deviation of 7.6 when the two outliers are excluded. The mean increased and the standard deviation decreased since these measures are affected by extreme observations. The median and quartiles should not change very drastically since these measures are not sensitive to extreme observations.

1.55

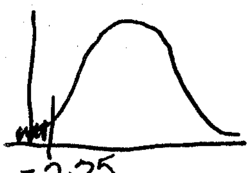
The average of $2.36 million is the mean. If it were the median, then half of the players would earn that salary.

1.72

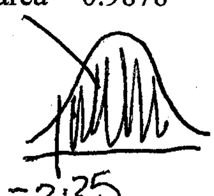
The curve with the higher peak has a standard deviation of approximately 0.2 and the shorter curve has a standard deviation of approximately 0.5

1.75

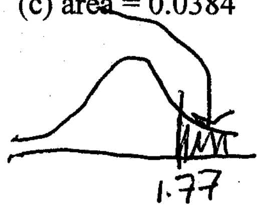
(a) area = 0.0122



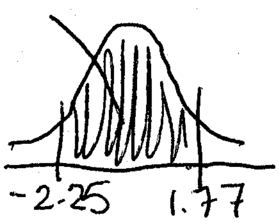
(b) area = 0.9878



(c) area = 0.0384



(d) area = .9494



1.76

(a) z = 0.84

(b) z = 0.39

1.79

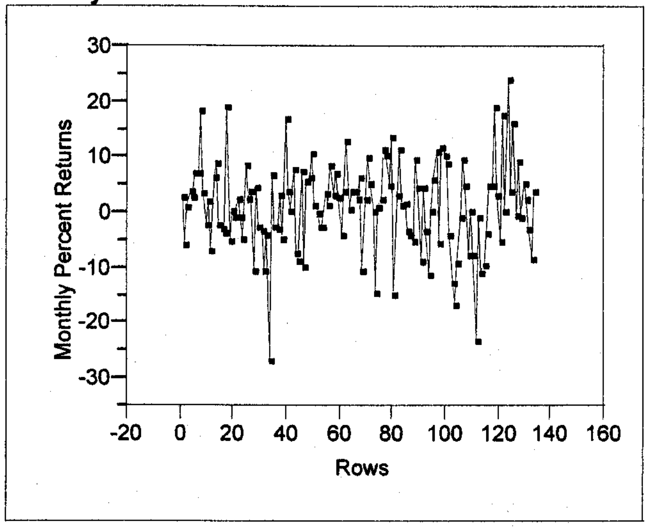
(a) –21% and 47%

(b) 22.36%

(c) 23.89%

1.90

Overlay Plot



This time plot shows that the returns did not display any obvious trends up or down over this 10 year period. The two outliers discovered in problem 1.53 are shown as the two low points on this plot. The two effects of action against smoking and sharp rises in stock prices may have produced this random-looking time plot.

1.91

Normal Corn

Quantiles

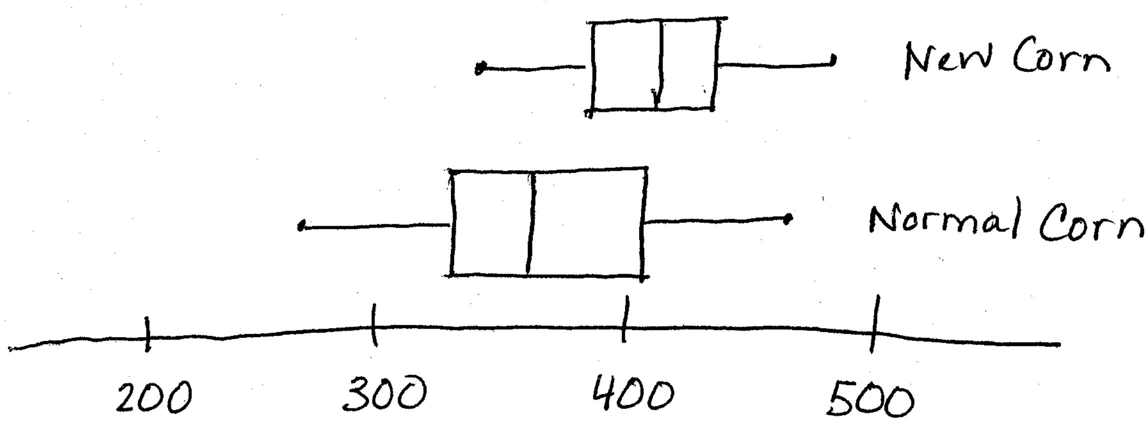
100.0%	maximum	462.00
99.5%		462.00
97.5%		462.00
90.0%		452.60
75.0%	quartile	401.25
50.0%	median	358.00
25.0%	quartile	333.00
10.0%		286.30
2.5%		272.00
0.5%		272.00
0.0%	minimum	272.00

New Corn

Quantiles

100.0%	maximum	477.00
99.5%		477.00
97.5%		477.00
90.0%		465.00
75.0%	quartile	429.25
50.0%	median	406.50
25.0%	quartile	379.25
10.0%		327.30
2.5%		318.00
0.5%		318.00
0.0%	minimum	318.00

(a) The JMP output above gives the five-number summaries for the new and normal corn.



From the box plots above, we can see that the new corn has a more symmetrical distribution than the normal corn with a smaller spread and higher median. There is a higher weight gain overall for new corn.

(b)

Normal Corn

Moments

Mean	366.3
Std Dev	50.805201
Std Err Mean	11.360388
upper 95% Mean	390.07757
lower 95% Mean	342.52243
N	20

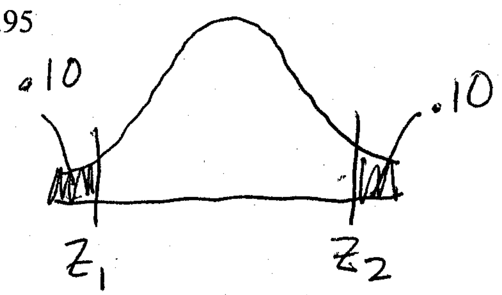
New Corn

Moments

Mean	402.95
Std Dev	42.728611
Std Err Mean	9.5544079
upper 95% Mean	422.94761
lower 95% Mean	382.95239
N	20

The mean weight gain of chicks fed the new corn is 36.65 grams higher than for chicks fed the normal corn.

1.95



$z_1 = -1.28$

$z_2 = 1.28$

$$z = \frac{x-\mu}{\sigma}$$

$$-1.28 = \frac{25-\mu}{\sigma}$$

$$1.28 = \frac{475-\mu}{\sigma}$$

$$-1.28\sigma = 25-\mu$$

$$\sigma = \frac{25-\mu}{-1.28} = \frac{\mu+25}{1.28}$$

$$\Rightarrow 1.28 = \frac{475-\mu}{\left(\frac{\mu+25}{1.28}\right)}$$

$$\mu+25 = 475-\mu$$

$$2\mu = 450$$

$$\mu = 250$$

$$\sigma = \frac{250+25}{1.28} = 175.78$$

$$\boxed{\begin{aligned}\mu &= 250\\ \sigma &= 175.78\end{aligned}}$$

3.9
The retail outlets could be labeled 1 – 440. The book suggests starting at line 105 which would produce the following SRS (simple random sample):
400, 077, 172, 417, 350, 131, 211, 273, 208, and 074.
Using JMP I obtained the following SRS:
86, 149, 307, 312, 171, 201, 314, 13, 88, 364.
Using JMP, answers will vary.

3.11
Mid-size: 001-500
Small Size: 0001-4400

You need to sample 5% of the mid-size accounts, therefore, the sample would be of size 25.
Sampling 1% of the small size would result in a sample of size 44.
The numbers obtained for the samples will vary.

4.3
0.105

4.7
(a) 0
(b) 1
(c) 0.01
(d) 0.6

4.14
0.54

4.16
(a) .45 – The sum of the probabilities must add to one
(b) 0.42

4.17
Model
I - illegitimate because the probabilities do not sum to one
II - legitimate
III – illegitimate because the probabilities do not sum to one
IV – illegitimate because probabilities cannot be greater than 1

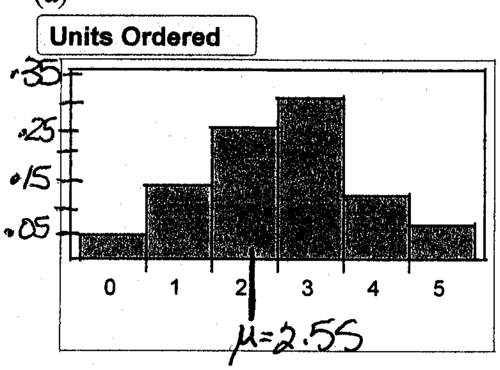
4.20
(a) 0.4
(b) 0.6
(c) 0.2
(d) 0.2

4.22
(a) $P(X \geq 32)$
(b) 0.025

4.30
(a) Abby & Deborah, Abby & Sam, Abby & Tonya, Abby & Roberto, Deborah & Sam, Deborah & Tonya, Deborah & Roberto, Sam & Tonya, Sam & Roberto, Tonya & Roberto
(b) 0.10
(c) 0.40
(d) 0.30

4.34
(a) all probabilities are between 0 and 1 and sum to one
(b) 0.94
(c) 0.86
(d) $P(X \geq 4) = 0.06$

4.44
(a)



(b) Mean = 2.55
(c) $P(X > 2.55) = 0.53$

4.46*
$\mu = 0.936\%$

4.47
$\sigma_Y^2 = 19225$
$\sigma_Y = 138.65$

4.67*
$\sigma_{X+Y} = 2796$
$\sigma_Z = 5{,}606{,}738$

4.99

(a) approximately 0

(b) We would need the shape of the distribution. In part (a) the CLT allowed us to approximate the probability associated with x-bar since n was "large"

4.108

(a) 0.3707

(b) mean of sample of 60: 100
  std error of sample of 60: 1.94

(c) 0.0049

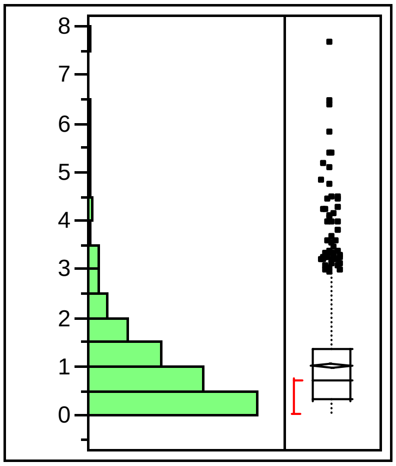
(d) only part (a) would be affected

LAB 1

Empirical Version of MMD&S Figure 4.17 (a)

Distributions

Column 1



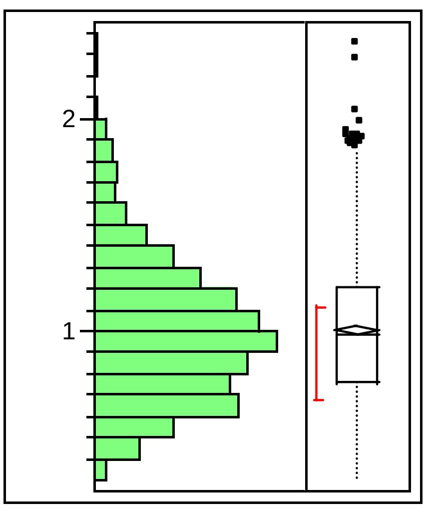
Moments

Mean	1.0000876
Std Dev	0.9881284
Std Err Mean	0.0312474
upper 95% Mean	1.0614056
lower 95% Mean	0.9387696
N	1000

Empirical Version of MMD&S Figure 4.17 (c)

Distributions

Column 11



Moments

Mean	1.0029826
Std Dev	0.3316185
Std Err Mean	0.0104867
upper 95% Mean	1.0235611
lower 95% Mean	0.9824041
N	1000