

Problem 1. Summary statistics show that variable w has sample mean 52 and sample variance 4. Provide the formula for a transformation of variable w that has sample mean 0 and sample variance 1.

Problem 2. Calculate $\sum_{i=1}^3 (2 + 6/i)$. Show all of your work.

Problem 3. For a sample of four thirty-year olds, the years of completed schooling are 12, 15, 13, and 12. Compute the sample mean, sample variance, and sample standard deviation. Is the data left- or right-skewed? Show all workings.

Problem 4. Suppose $X = 10$ with probability 0.6, $X = 20$ with probability 0.3, and $X = 30$ with probability 0.1. Compute the mean, variance, and standard deviation of X . Show your work.

Problem 5. Suppose for $X \sim (200, 10^2)$, we form 100 samples of size 25 and obtain 100 sample means \bar{x} . What approximately do you expect the average of the \bar{x} to equal? What about the standard deviation of \bar{x} ? Explain your answer.

Problem 6. Consider the following Stata output.

```
. sum highmeaning, d
```

Fraction who say their work makes the world a
better place

	Percentiles	Smallest		
1%	.28	.26		
5%	.41	.28		
10%	.45	.33	Obs	194
25%	.54	.34	Sum of Wgt.	194
50%	.68		Mean	.6662887
		Largest	Std. Dev.	.1602189
75%	.8	.93		
90%	.88	.93	Variance	.0256701
95%	.91	.95	Skewness	-.1739289
99%	.95	.97	Kurtosis	2.164972


```
. sum major earlycareer midcareer highmeaning
```

Variable	Obs	Mean	Std. Dev.	Min	Max
major	0				
earlycareer	194	56257.22	14535.76	36100	139000
midcareer	194	85715.46	22338.64	48100	173000
highmeaning	194	.6662887	.1602189	.26	.97


```
t_193, 0.005 = 2.6015
t_193, 0.010 = 2.346
t_193, 0.025 = 1.9723
t_193, 0.050 = 1.6528
t_193, 0.100 = 1.3062
```

- (a) Does variable `highmeaning` appear to be symmetrically distributed? Explain.
- (b) If variable `highmeaning` was normally distributed, then what range of values would you expect 95% of observations to lie in? Explain.

- (c) Provide a 90 percent confidence interval for the population mean earnings for Masters graduates with 0-5 years of work experience. Show your work.
- (d) What Stata command would enable you to directly answer part (c)?
- (e) The claim is made that population mean earnings for Masters graduates with 0-5 years of work experience equals \$60,000. Test this claim at significance level 0.10. State clearly the null and alternative hypothesis and your conclusion.
- (f) Suppose we perform a test of a hypothesis and find that $p = 0.06$. Will we reject or not reject the null hypothesis at significance level 5%? Explain your answer.

Problem 7. Consider the following Stata output

```
. summarize growth, detail
```

Percentage annual growth in variable real				
	Percentiles	Smallest		
1%	-.5972697	-.5972697		
5%	.4553079	.4553079		
10%	1.010702	1.010702	Obs	28
25%	2.056625	1.037691	Sum of wgt.	28
50%	3.288075		Mean	4.155236
		Largest	Std. Dev.	2.974328
75%	5.993356	8.543264		
90%	8.990582	8.990582	Variance	8.846628
95%	10.00192	10.00192	Skewness	.6414501
99%	10.44776	10.44776	Kurtosis	2.458796

Calculate the interquartile range of variable `growth`.

Problem 8. Consider the following Stata output.

```
. sum diff, detail
```

diff				
	Percentiles	Smallest		
1%	-100.7092	-107.1325		
5%	-36.7719	-104.92		
10%	-20.9986	-100.7092	Obs	274
25%	-8.998913	-61.88319	Sum of wgt.	274
50%	-2.2891		Mean	1.421385
		Largest	Std. Dev.	35.51708
75%	2.870001	162.9905		
90%	19.2143	175.58	Variance	1261.463
95%	70.69516	180	Skewness	2.505035
99%	175.58	215.399	Kurtosis	14.55209

Does variable `diff` appear to be normally distributed? Explain.

Problem 9. An estimator is the best among consistent estimators if

- (a) it is also unbiased,
- (b) it has the smallest variance,
- (c) it is normally distributed,
- (d) none of the above.