Homework 5 11.1 # 9b

a) $\alpha = 0.05$

	n	$\bar{x}$	s
$X_1 =$	45	55	8
$X_2 =$	45	60	3

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$df = 44$

$CV_{\frac{\alpha}{2}} = -2.015$

$$\frac{55 - 60}{\sqrt{\frac{8^2}{45} + \frac{3^2}{45}}} = -3.926$$

$-3.926 < -2.015$

$t < CV$ Reject Null

b) A the $\alpha = 0.05$ level of confidence there is sufficient evidence to reject the claim that the average amounts charged are different.

11.1 prob 11

a)

$n_1 = 35 \qquad \bar{x}_1 = 700 \qquad s_1 = 30$

$n_2 = 35 \qquad \bar{x}_2 = 710 \qquad s_2 = 35$

$$\frac{700-710}{\sqrt{\frac{30^2}{35}+\frac{35^2}{35}}} = -1.283$$

$df = 34$

$\alpha = 0.10 \qquad CV_\alpha = -1.307$

$-1.283 > -1.307$

$t > CV_\alpha$

Fail to reject

For $\alpha = 0.10$ there is not sufficient evidence to reject the claim that time until failure is the same.

b) $p = 0.1003$

C) $\alpha = 0.10$

$p > \alpha$

$0.1003 \quad 0.05$

No

11.2 # 6

a) $\alpha = 0.025$, $n_1 = 13$, $n_2 = 25$ Left

$df = n_1 + n_2 - 2$

$df = 13 + 25 - 2 = 36$

$CV = -2.028$

b) $\alpha = 0.005$, $n_1 = 7$, $n_2 = 18$, Right

$df = 7 + 18 - 2 = 23$

$CV = 2.819$

c) $\alpha = 0.1$, $n_1 = 15$, $n_2 = 15$, two-tailed

$\frac{\alpha}{2} = 0.05$

$df = 15 + 15 - 2 = 28$

$CV = (-1.701, 1.701)$

11.2 # 86

	n	$\bar{x}$	s
x_1	16	5	1
x_2	15	6	2

$df = 29$

d) $\alpha = 0.10$ Left $CV = -1.311$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Equal Variances

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\frac{(16-1)1^2 + (15-1)2^2}{16+15-2} = 2.45 \qquad \frac{5-6}{\sqrt{2.45\left(\frac{1}{15} + \frac{1}{16}\right)}} = -1.778$$

b) Yes, Reject H_0

11.2 # 10 b c

b)

	n	$\bar{x}$	s
X_1	15	180	50
X_2	13	150	30

$H_0: X_1 = X_2$

$X_1 > X_2$

Right

$$s_p^2 = \frac{(15-1)50^2 + (13-1)30^2}{15+13-2} = 1761.538$$

$$\frac{15-13}{\sqrt{1761.5\left(\frac{1}{15}+\frac{1}{13}\right)}} = 1.886 = t$$

$\alpha = 0.01$ $df = 26$

C.V. = 2.479 Fail to reject

11.3 #8

a) Yes, the book is the same for each price.

```python
import pandas as pd
import seaborn as sns

df = pd.DataFrame({
    'bookstore' : [70,38,88,165,80,103,42,98,89,97,140,40,175,85,100,68,67,140,49,149,126,92,144,98,40],
    'onlineRetailer' : [60,36,89,149,136,95,50,111,65,86,130,30,150,75,85,62,69,142,40,127,130,93,129,84,52]
})
```

b) That the textbook price difference follows a normal distribution

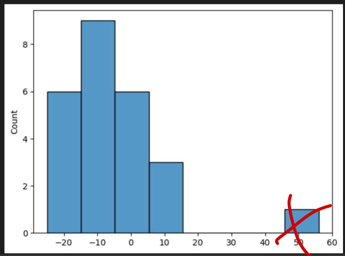
c) Ignoring the outlier the histogram of the differences series appears to be normal.

```python
priceDifference = []
for _,row in df.iterrows():
    priceDifference.append(row.iloc[1] - row.iloc[0])

differenceSeries = pd.Series(priceDifference)
diffMean = differenceSeries.mean()
diffStd = differenceSeries.std()
sns.histplot(differenceSeries)
```

✓ 0.0s

<Axes: ylabel='Count'>



e) $\alpha = 0.01$ $df = 24$

CV = -2.492

```python
import scipy.stats as stats
stats.ttest_1samp(differenceSeries, 0).confidence_interval(0.99)
```

[11] ✓ 0.0s

```
ConfidenceInterval(low=-13.430068915746416, high=4.790068915746415)
```

Fail to reject

```python
t, p = stats.ttest_1samp(differenceSeries, 0)
print("t: " + str(t) + "\nmean " + str(diffMean) + "\nstandard deviation: "+ str(diffStd))
```

✓ 0.0s

```
t: -1.3263103465368828
mean -4.32
standard deviation: 16.285781119328195
```

At $\alpha = 0.01$ there is not sufficient evidence to suggest that online prices are lower than in the bookstore

11.3 # 9 abe

a) Yes, because the cashier is the same between samples, only the register changes

b) That the distribution of the differences is approximately normal.

e) $\alpha = 0.05$

```python
df = pd.DataFrame({
    'oldCashRegister' : [60,70,55,75,62,52,58],
    'newCashRegister' : [65,71,55,75,65,57,57]
})
priceDifference = []
for _,row in df.iterrows():
    priceDifference.append(row.iloc[1] - row.iloc[0])
stats.ttest_1samp(priceDifference, 0, alternative="greater")
```

✓ 0.0s

```
TtestResult(statistic=1.9824814143238605, pvalue=0.04734883986144706, df=6)
```

$p < \alpha$

$0.047 \quad 0.05$

Fail to reject

$t = 1.98 \quad p = 0.047 \quad df = 6$

At $\alpha = 0.05$ there is sufficient evidence to support the claim that the new cash register increases the amount of items the cashier can process.

11.3 # 10 abe

```python
df = pd.DataFrame({
    'modelA' : [150,145,160,155,152,153],
    'modelB' : [152,146,160,157,154,155]
})
priceDifference = []
for _,row in df.iterrows():
    priceDifference.append(row.iloc[1] - row.iloc[0])

stats.ttest_1samp(priceDifference, 0)
```

[23] ✓ 0.0s

```
TtestResult(statistic=4.391550328268399, pvalue=0.007077597925498281, df=5)
```

$t = 4.39$ $p = 0.007$ $df = 5$

a) paired design is appropriate because the driver doesn't change between models

b) That the distribution of the difference series is approximately normal.

e) $\alpha = 0.10$

$$\underset{p}{0.007} < \underset{\alpha/2}{0.05}$$ Reject Null

At $\alpha = 0.10$ there is sufficient evidence to suggest that there is a difference in the braking difference between the two models.

11.4 # 10

$\hat{p}_1 = 0.64 \qquad n_1 = 300$

$\hat{p}_2 = 0.63 \qquad n_2 = 250$

$\alpha = 0.05$

$\alpha/2 = 0.025$

t Statistic

$$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Value

$$\frac{0.64 - 0.63}{\sqrt{0.66\,(1-0.66)\left(\frac{1}{300}+\frac{1}{250}\right)}} = 1.478$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n}$$

$$.66 = \frac{0.64(300) + 0.63(250)}{300 + 250} = \hat{p}$$

Fail to reject

$CV = 1.960$

11.4 # 11

	n	x
1	1000	450
2	〃	655

a) $\alpha = 0.01$ right tail $\hat{p}_1 = .45$

$CV = 2.326$ $\hat{p}_2 = .655$

$$\hat{p} = \frac{450 + (655)}{2000} = 0.55$$

$$t = \frac{.655 - .45}{\sqrt{0.55(0.45)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 9.214$$

↓

Reject Null