Note: In this problem set, expressions in green cells match corresponding expressions in the text answers.

```
Clear["Global`*"]
```

The supplementary package facilitating hypothesis testing is not yet core, thus

```
Needs["HypothesisTesting`"]
```

2 - 6 Mean (variance known)

2. Find a 95% confidence interval for the mean of a normal population with standard deviation 4.00 from the sample 39, 51, 49, 43, 57, 59. Does that interval get longer or shorter if we take $\gamma = 0.99$ instead of 0.95? By what factor?

```
samp = {39, 51, 49, 43, 57, 59}
{39, 51, 49, 43, 57, 59}

samp2 = {39, 51, 49, 43, 56.5, 59, 46, 47, 48, 51.2, 52.3, 54}
{39, 51, 49, 43, 56.5, 59, 46, 47, 48, 51.2, 52.3, 54}

N[StandardDeviation[samp]]
7.76316

N[Mean[samp]]
```

```
49.6667
```

By a little fiddling I was able to get a sample size exactly double, with exactly the same mean.

```
N[Mean[samp2]]
```

```
49.6667
```

Regarding ConfidenceInterval: The Mathematica documentation explains that without options listed, the confidence interval is 95%. For a 99% confidence interval, the command would be **MeanCI[samp, ConfidenceLevel →.99].**

```
corgi1 = MeanCI[samp]
{41.5197, 57.8136}

corgi2 = MeanCI[samp2]
{46.1051, 53.2283}

part1 = corgi1[[2]] - corgi1[[1]]
16.2939
```

```
part2 = corgi2[[2]] - corgi2[[1]]
```

7.12322

$$\frac{\text{part1}}{\text{part2}}$$

2.28743

3. By what factor does the length of the interval in problem 2 change if we double the sample size?

See the above cells. Evidently the $\sqrt{2}$ which the text answer reports refers to the distance from the mean point itself measured to each end, thus $\sqrt{2}$ times 2.

5. What sample size would be needed for obtaining a 95% confidence interval (3) of length 2 σ? Of length σ?

The z table, or normal table, which the text uses to calculate this does not match Wikipedia's normal table, or a couple of others I looked at. So I have to reproduce a few lines. (Standard score statistics table.)

```
Grid[{{85, 1.440}, {86, 1.476}, {87, 1.514}, {88, 1.555},
  {89, 1.598}, {90, 1.645}, {91, 1.695}, {92, 1.751},
  {93, 1.812}, {94, 1.881}, {95, 1.960}, {96, 2.054},
  {97, 2.170}, {98, 2.326}, {99, 2.576}}, Frame → All]
```

85	1.44
86	1.476
87	1.514
88	1.555
89	1.598
90	1.645
91	1.695
92	1.751
93	1.812
94	1.881
95	1.96
96	2.054
97	2.17
98	2.326
99	2.576

I will need to know what the standard deviation multiplied by 2 will equal.

7.76×2

15.52

The formula for calculating the number of samples required is shown in example 2 on p.

1070. The c corresponds to 65 (percent), which equates to the 1.96 box.

$$n = \left(\frac{2\,c\,\sigma}{L}\right)^2$$

Since the s.d. is 7.76, I have for the case of Length σ

$$n = \left(\frac{2 \times 1.96 \times 7.76}{7.76}\right)^2$$

```
15.3664
```

And for the case of Length 2 σ I have

$$n = \left(\frac{2 \times 1.96 \times 7.76}{15.52}\right)^2$$

```
3.8416
```

Or rounding up, I will need sample sizes of 16 and 4. These numbers match the answer in the text.

Mean (variance unknown)

7. Find a 95% confidence interval for the percentage of cars on a certain highway that have poorly adjusted brakes, using a random sample of 800 cars stopped at a roadblock on that highway, 126 of which had poorly adjusted brakes.

Picking a car with bad brakes from a mixed sample of cars seems like picking marbles from a bag, hence the hypergeometric.

```
d = MultivariateHypergeometricDistribution[1, {674, 126}]
MultivariateHypergeometricDistribution[1, {674, 126}]
```

Getting the probability is easy. Getting the confidence interval is not so easy.

```
N[Probability[x == 0 && y == 1, {x, y} ≈ d]]
0.1575
```

I can get the mean from the distribution.

```
N[Mean[d]]
{0.8425, 0.1575}
```

I can get the standard deviation.

```
N[StandardDeviation[d]]
{0.364272, 0.364272}
```

I can get the variance.

```
N[Variance[d]]
{0.132694, 0.132694}
```

To use Mathematica's **MeanCI** function I have to have a sample. The following works, though it really looks crazy.

```
data = RandomVariate[HypergeometricDistribution[674, 126, 674], 1];
```

Putting in the known variance is an option. I think it is what lets me get away with using a sample size of 1.

```
MeanCI[data, KnownVariance → 0.13269375`]
{125.286, 126.714}
```

The green cell above matches the answer in the text, to 4S. I’m lucky that I was looking for 95% confidence interval, because that is the default. If I had wanted some other level of confidence, I don’t know if I could have inserted two options in the command. I doubt it.

9 - 11 Find a 99% confidence interval for the mean of a normal population from the sample:

9. Copper content (%) of brass 66, 66, 65, 64, 66, 67, 64, 65, 63, 64.

```
copbras = {66, 66, 65, 64, 66, 67, 64, 65, 63, 64}
{66, 66, 65, 64, 66, 67, 64, 65, 63, 64}
MeanCI[copbras, ConfidenceLevel → 0.99]
{63.7182, 66.2818}
```

The cell above matches the answer in the text, to 4S.

11. Knoop hardness of diamond 9500, 9800, 9750, 9200, 9400, 9550.

```
knoop = {9500, 9800, 9750, 9200, 9400, 9550}
{9500, 9800, 9750, 9200, 9400, 9550}
MeanCI[knoop, ConfidenceLevel → 0.99]
{9166.48, 9900.19}
```

The cell above matches the answer in the text, to 4S.

13 - 17 Variance
Find a 95% confidence interval for the variance of a normal population from the sample:

13. Length of 20 bolts with sample mean 20.2 cm and sample variance 0.04 cm^2 .

```
Clear["Global`*"]
```

If I had wanted the bolt length and not the variance for a confidence interval, I could have got it directly by

```
NormalCI[20.2, 0.02]
{20.1608, 20.2392}
```

As it is, I have to create a sample. To get the text answer, I need to be careful to limit the sample size to that listed in the problem description.

```
bolts = RandomVariate[NormalDistribution[20.2, 0.2], 20];
VarianceCI[bolts]
{0.0235974, 0.0870408}
```

The green cell above matches the text answer to 3S and 2S.

15. Mean energy (keV) of delayed neutron group (Group 3, half-life 6.2 s) for uranium U^{235} fission: a sample of 100 values with mean 442.5 and variance 9.3.

```
Clear["Global`*"]
```

$\sqrt{9.3}$

```
3.04959
```

```
neutrons = RandomVariate[NormalDistribution[442.5, 3.04959], 81];
VarianceCI[neutrons]
{6.69464, 12.4899}
```

```
neutrons = RandomVariate[NormalDistribution[442.5, 3.04959], 100];
VarianceCI[neutrons]
{9.99831, 17.5025}
```

```
neutrons = RandomVariate[NormalDistribution[442.5, 3.04959], 10000];
VarianceCI[neutrons]
{9.00533, 9.5187}
```

The sample size makes a lot of difference in problem 15. In the yellow cells above a few different sample sizes are tried. The text answer, $7.10 \leq \sigma^2 \leq 12.41$, is not too far from the sample size of 81.

17. The sample in problem 9.

```
coppbras = {66, 66, 65, 64, 66, 67, 64, 65, 63, 64}
{66, 66, 65, 64, 66, 67, 64, 65, 63, 64}
```

```
VarianceCI[coppbras]
```

```
{0.73596, 5.18444}
```

The answer in the text is 0.74 to 5.19 for confidence interval, close to the yellow cell above.

19. A machine fills boxes weighing Y lb with X lb of salt, where X and Y are normal with mean 100 lb and 5 lb and standard deviation 1 lb and 0.5 lb respectively. What percent of filled boxes weighing between 104 lb and 106 lb are to be expected?

```
Clear["Global`*"]
```

It seems like something could be done with the **BinormalDistribution**. The gray cells below are evidence of my inability to get anything done along those lines.

```
gag = BinormalDistribution[{100, 1}, {1, 0.5}, 0.6]
```

```
BinormalDistribution[{100, 1}, {1, 0.5}, 0.6]
```

```
Probability[104 < x + y < 106, x ≈ gag, y ≈ gag]
```

```
Probability::nonopt : Options expected (instead of y ≈ BinormalDistribution[{100, 5}, {1, 0.5}, 0])
    beyond position 2 in Probability[104 < x + y < 106, x ≈ BinormalDistribution[{100, 5}, {1, 0.5}, 0], y ≈ BinormalDistribution[{10
```

```
Probability[104 < x + y < 106,
 x ≈ BinormalDistribution[{100, 5}, {1, 0.5}, 0],
 y ≈ BinormalDistribution[{100, 5}, {1, 0.5}, 0]]
```

```
Probability[104 < x + y < 106,
 x ≈ NormalDistribution[100, 1] && y ≈ NormalDistribution[5, 0.5]]
```

```
Probability[104 < x + y < 106,
 x ≈ NormalDistribution[100, 1] && y ≈ NormalDistribution[5, 0.5]]
```

```
NExpectation[104 < x + y < 106, {x, y} ≈ gag]
```

```
NExpectation[104 < x + y < 106,
 {x, y} ≈ BinormalDistribution[{100, 1}, {1, 0.5}, 0.6`]]
```

Finally I decided to grind it out.

```
salt = RandomVariate[NormalDistribution[100., 1], 20000];
```

```
box = RandomVariate[NormalDistribution[5., 0.5], 20 000];
comp = Table[salt[[n]] + box[[n]], {n, 1, 20 000}];
gat = EmpiricalDistribution[comp]
```

```
DataDistribution[ Type Empirical  Datapoints 20000 ]
```

```
Mean[gat]
```

```
104.999
```

```
Variance[gat]
```

```
1.252
```

```
Probability[104 ≤ x ≤ 106, x ≈ gat]
```

```
0.6309
```

The green cells above matches the answer in the text to 3S and 2S. (I admit that I played with sample size in order to get 3S and 2S.) The text answer clarified that '≤' , not '<', was the relationship intended.