Supplementary exercise 7.64 of IPS7e

(same data set as Supplementary Exercise 6.95)

Data: 12 readings, $X_1, \ldots, X_{12}$, of home radon detectors when exposed to 105 picocuries per liter of radon; the purpose being to examine the accuracy of the detectors.

Model: $X_1, \ldots, X_{12}$ are i.i.d. (a simple random sample) from a distribution with unknown mean, median and standard deviation. Note that for this analysis we do not assume a known value of the standard deviation.

(a) Some descriptive statistics from Minitab (including the stemplot):

Descriptive Statistics: radon

Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Skewness
radon	12	0	104.13	2.71	9.40	91.90	96.90	102.75	109.90	122.30	0.85

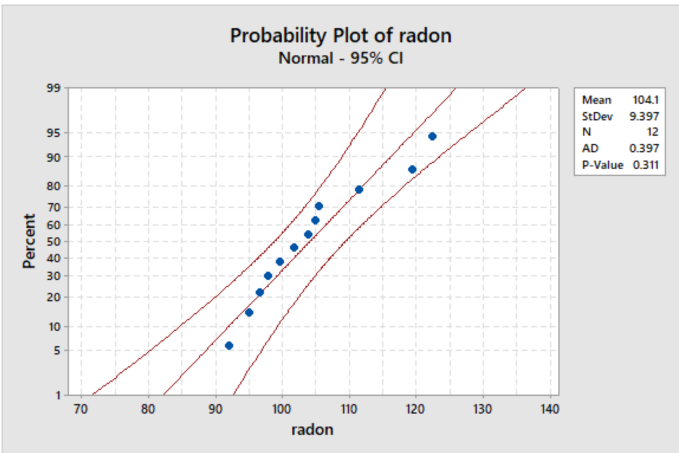
Variable	Kurtosis
radon	-0.01

Stem-and-Leaf Display: radon

Stem-and-leaf of radon N = 12

```
1    9   1
5    9   5679
(3)  10  134
4    10  5
3    11  1
2    11  9
1    12  2

Leaf Unit = 1
```



Comments:
The listings shows the values for the sample mean ($\hat{\mu} = \bar{X} = 104.1$), the sample median (102.8) and the sample standard deviation ($\hat{\sigma} = s = 9.40$). The stemplot shows (so would a dotplot) that the distribution is somewhat right-skewed (skewness=0.85), but the normality test is far from significant. There are no obvious outliers.

The statement in the problem text that the skewness is "not strong enough to forbid use of the t procedures" is maybe a bit surprising when comparing to textbook guidelines (PSLS 3e p.

426; IPS 7e p. 417-18). Because the sample size is small (< 15), we should consider whether the data are close to normal; that seems a bit questionable with the skewness. It is worth recalling that a non-significant normality test is no proof that the data are truly normally distributed, it just tells us that there is not enough evidence to say they are not normally distributed. With the fairly small sample size, there is not much power to (statistically) detect deviations from normality. On the other hand, the guidelines state that we should not use t procedures if the data are clearly non-normal or if outliers are present, and none of these two cases apply here.

Note also that the assumption of normality is substantially more important when using t procedures than z procedures (for when the standard deviation is assumed known). This is because estimation of the standard deviation from the data can be substantially affected by non-normality, so that the t-distribution will no longer apply as a reference distribution. We will carry out an alternative analysis using non-parametric methods in a later exercise.

(b) We set up the hypotheses as

$$H_0 : \mu = 105 \quad \text{versus} \quad H_a : \mu \neq 105.$$

We chose a two-sided alternative hypothesis because there is no indication of any interest in a specific direction; in fact, the question referred to whether the mean *differed* from the true value. We will use the Minitab menu under `Basic Statistics-1 Sample t` for the calculation.

One-Sample T: radon

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
12	104.13	9.40	2.71	(98.16, 110.10)

μ: mean of radon

Test

Null hypothesis H₀: μ = 105
Alternative hypothesis H₁: μ ≠ 105

T-Value	P-Value
-0.32	0.755

Comments:
The t-test is clearly non-significant at $t = -0.32$ and a P-value of 0.76. There is absolutely no evidence that the reading of the detectors differ systematically from the true value of 105 picocuries per liter. That is good but does not by itself mean that the results obtained are satisfactory. One would certainly also want to consider the standard deviation in the readings which seems quite large (and the actual values scatter considerably around the true value of 105). So we might be in a "true mean, large scatter" situation (compare the figure on slide 5L-2). Even the 95% confidence interval for μ is not very narrow around 105, and the measurements themselves are much more variable than the mean (as we saw in Exercise 6.33).

It is perhaps interesting to compare the results with those of Exercise 6.95 where the standard deviation was assumed to be known (and equal to 9). Generally speaking, the analysis with unknown standard deviation is weaker and should lead to larger confidence intervals and higher P-values (because the t distribution is wider than the standard normal). The 95% CI is indeed about 2 units wider, and this is also due to the sample standard deviation being a bit above the assumed value of 9. The P-values are however quite similar, and this is because the difference between the normal and $t(11)$ distributions shows mostly in the tails whereas the observed values (around -0.3) are still in the centre of the distributions.