

Feedback – Lecture 03A: univariate data analysis

Important note: As mentioned in class last Friday, there were several requests to post the quizzes earlier.

To help you work on the quiz earlier, the quizzes will be based on the textbook material.

This quiz is based on material from pages 38 to middle of 45 from the course textbook. Videos will be posted on this content on Monday.

You have 1 attempt for the quiz. Please read the instructions carefully. Please double check your answers before submitting.

Solutions will be released when the quiz closes, at 09:25am, Tuesday, 20 January 2015.

Question 1

Select *all features* that apply to the normal distribution:

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> The approximate area between $-\sigma$ and σ is 70%	✓ 0.25	
<input checked="" type="checkbox"/> The standard deviation is indicated with the symbol σ .	✓ 0.25	
<input checked="" type="checkbox"/> It is symmetric in appearance.	✓ 0.25	
<input checked="" type="checkbox"/> The peak value is usually indicated with the symbol μ , representing the mean.	✓ 0.25	
Total	1.00 / 1.00	

Question 2

Have you installed and used the `car` library in R. (you must answer "Yes" to get full grade). If you

haven't yet done so, please do so, by following the [software tutorial](#), steps 11, 12 and 13.

Your Answer	Score	Explanation
<input checked="" type="radio"/> Yes, I have installed the <code>car</code> library and tried to use the <code>qqPlot(...)</code> function.	✓ 1.00	
<input type="radio"/> No, not yet.		
Total	1.00 / 1.00	

Question 3

For the standardized normal distribution, for a variable $z \sim \mathcal{N}(0, 1)$

(check all that correctly apply)

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> In R, writing <code>pnorm(1, mean=0, sd=1)</code> is the same as <code>pnorm(1)</code>	✓ 0.20	
<input checked="" type="checkbox"/> In R, using <code>pnorm(0)</code> gives a value of 0.0 as well.	✗ 0.00	
<input checked="" type="checkbox"/> if there was a probability of 2.5% of observing something, that would correspond to a z value of -1.96 .	✓ 0.20	
<input checked="" type="checkbox"/> as $z \rightarrow \infty$, the cumulative area under the curve also $\rightarrow \infty$.	✗ 0.00	The distribution is finite, and the cumulative area $\rightarrow 1.0$
<input checked="" type="checkbox"/> the probability of observing a value of $z \leq -1$ is about 16%.	✓ 0.20	
Total	0.60 / 1.00	

Question 4

A food production facility fills bags with muffin mix with a listed bag weight of 500 grams. The packaging system is set to fill bags with a mean weight of 520 grams. Long term data from the facility shows the variance of fill weights is 3.16 grams, and a q-q plot of the fill weights confirms a normal distribution.

Out of 5000 boxes, approximately how many will be under the specified weight of 500 grams?

Your Answer	Score	Explanation
<input type="radio"/> 13		
<input type="radio"/> 57		
<input type="radio"/> 115		
<input type="radio"/> 274		
<input checked="" type="radio"/> None	✓ 2.00	
Total	2.00 / 2.00	

Question Explanation

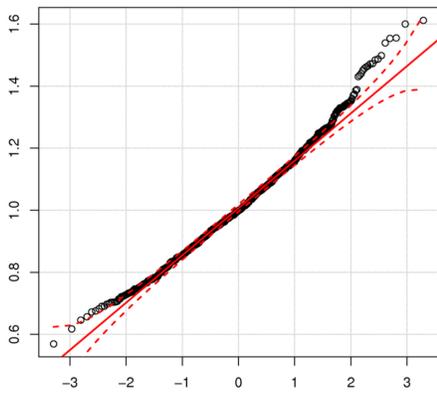
None: this small standard deviation, $s = 1.78 = \sqrt{3.16}$ grams, means essentially all the boxes are filled with around 520 grams of product. Note: the question was graded correctly, as stated. My intention, however was to ask for the number of under-filled boxes, using a **standard deviation** of 3.16 grams. Using that revised value, can you show that the number of under filled boxes would be around 115?

Question 5

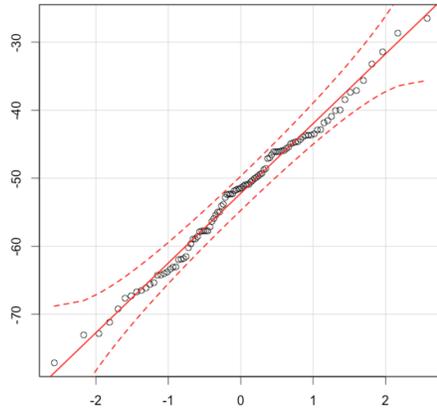
Which of these data shown are from a normal distribution? The q-q plots have the normal z on the x-axis in all cases, and 95% bounds are shown.

(check all that apply)

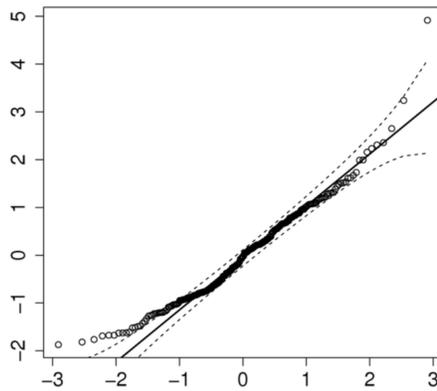
Your Answer	Score	Explanation
<input checked="" type="checkbox"/>	✗ 0.00	This distribution has a heavy right tail.



✓ 0.25

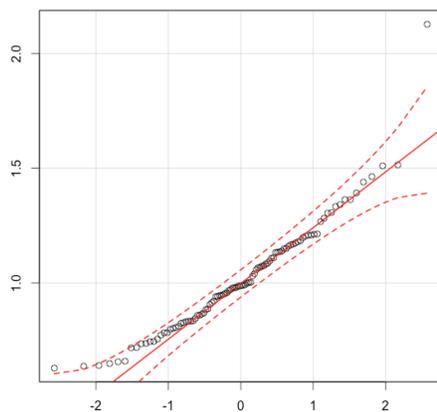


✗ 0.00



✓ 0.25

A few outliers are expected; remember these are 95% bounds, so 1 in 20 data points outside the limits can be expected. In this case, there is one extreme outlier at the



Total

0.50 /
1.00

Question 6

The data of the temperatures from your company's baking ovens in Hamilton follow an approximate [F-distribution](#), and range between 100 and 150°C.

You want to compare the data to another baking facility in Vancouver, so you start by standardizing them. After standardization:

(check all that correctly apply)

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> the variance of the data set will be 1.0	✓ 0.25	Yes, so is the standard deviation.
<input checked="" type="checkbox"/> the mean of the standardized data will be zero.	✓ 0.25	
<input checked="" type="checkbox"/> the median of the standardized data will be zero.	✗ 0.00	No; you can't say anything about the median. Try it with a data set; the median will just shift, but not necessarily to zero.
<input checked="" type="checkbox"/> the standard deviation of the data set will be 1.0	✓ 0.25	Yes, so is the variance.
Total	0.75 / 1.00	