Exam pincode 25619

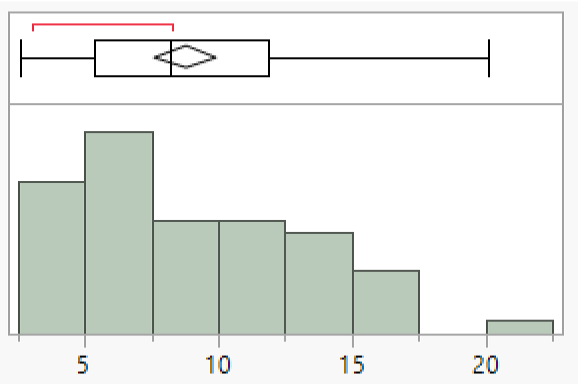
NB: All P-values are calculated

Problem 1

1. P(Rhesus positive)=84%
2. P(Rhesus positive given Antigens)=58.333%
3. P(Rhesus positive and antigens in at least 3 out of five)=48.1255%

Problem 2

1.



The distribution looks unimodal with a lot of right skew. This can also be seen from the boxplot with the right antenna being bigger much bigger than the left.

The empirical rule states that

Approximately 68% of the observations is in the interval]4.3109;13.2264[

23 of the observations is outside (61.667% is inside the interval)

Approximately 95% of the observations is in the interval]-0,1469;17.6842[

1 is outside (98,33% is inside)

Almost all of the observations is in the interval]-4.6046;22.1419[None is outside (100% is inside)

The empirical rule is fairly accurate on these data, although more of the observations than supposed is outside 1 standard deviation of the mean. With more data this would be more accurate due to the law of large numbers.

Problem 3

1. The Z-statistic is -73,88998.

Approximately normally distributed under the null hypothesis this gives a P-value below 0.0001. We can therefore reject the null hypothesis and conclude that the proportion of women in default on their student loans differs from the number of women taking student loans. With the proportion of women on default being approximately 43% we can conclude that more women than others repay their student loans

Problem 4

Student Number: 19-10-2021
Exam name: Statistics (BINBO1139E+BISHO1005E)

1. The 95% confidence interval of a man being hired compared to a woman is approximately normally distributed under the null hypothesis and has the interval]-0.15607;-0.041597[. On those grounds we can conclude that women are more likely to get the job compared to a man, because the 95% confidence interval doesn't include 0. This means that we with 95% confidence can say that women are between 4.16% and 15.61% more likely to get the job than men.

Problem 5

1.

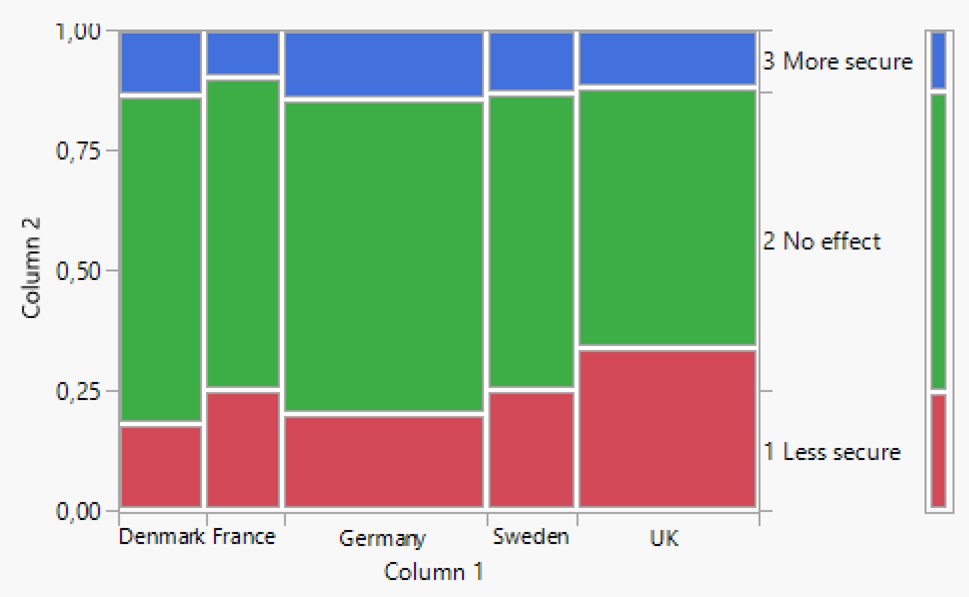
Tests

N	DF	-LogLike	RSquare (U)
4978	8	50,547362	0,0112

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	101,095	<,0001*
Pearson	102,373	<,0001*

The Pearson X^2-test statistic is 102.373, which approximately X^2-distributed under the null hypothesis with 8 degrees of freedom gives a P-value below 0.0001. We can thereby conclude that job security during the corona virus outbreak differs between the countries.

Mosaic Plot



As can be seen in the Mosaic plot a bigger proportion have less job security in the UK than in Denmark and Germany, with France and Sweden being in between. A bigger proportion felt more secure in their jobs in Germany, Denmark and Sweden with France having the smallest proportion.

Problem 6

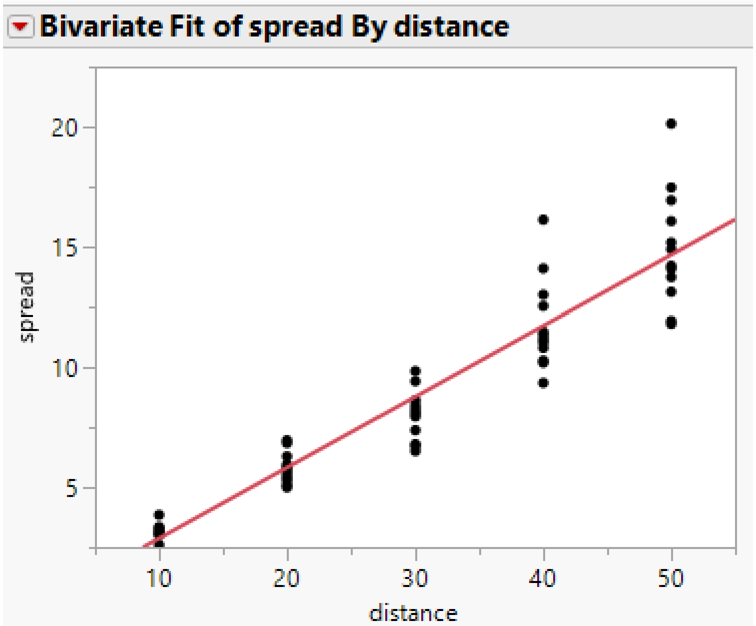
1.

The 90% confidence interval is based on a Two sampled T-test assuming the same standard deviation and the T-quantile with 58 degrees of freedom. The confidence interval is

]-1.25064;2.61864[

This means that we cannot conclude a difference in spread between the two types of cartridges. We can with 90% confidence conclude that the type B cartridge spreads between 1.25 inches less and 2.61864 inches more than the type A cartridge.

2.

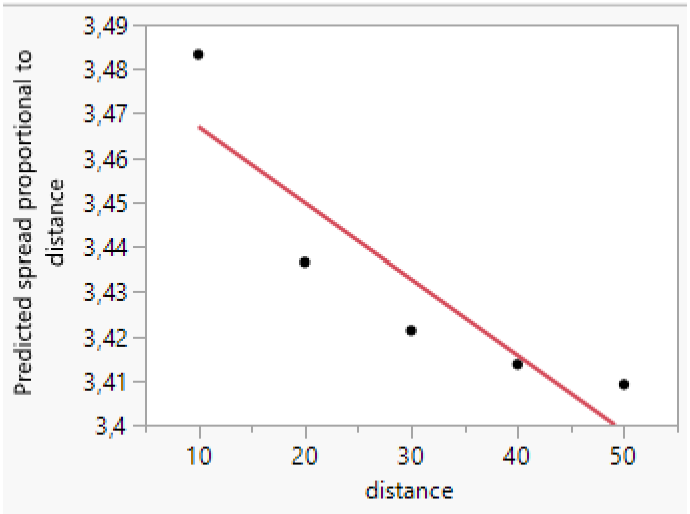


Linear Fit

Linear Fit

spread = -0,078083 + 0,2948917*distance

The regression equation is Spread(y)=-0.078083+0.2948917*distance(x).



As can be seen from the graph there seems to be a nonlinear correlation between the proportions. From a short distance the expected spread is proportionally bigger to the distance than from longer distances, but the effect seems to fade the longer you go away.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	0,03519580	0,035196	246,0306
Error	58	0,00829716	0,000143	**Prob > F**
C. Total	59	0,04349296		<,0001*

The ANOVA gives an F-Value of 246.0306, which approximately F-distributed under the null hypothesis with 1 and 58 degrees of freedom gives a P-value below 0.0001. We can thereby conclude, that the proportion of the spread differs at different distances, and the spread is not proportional to the distance to the target.

3.

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
type	1	1	7,0178	3,2470	0,0769
distance	1	1	1043,5331	482,8233	<,0001*
type*distance	1	1	0,8417	0,3894	0,5351

The F-test statistic for the dependence of type on the effect of distance on the spread is 0.3894, which approximately F-distributed under the null with 1 and 56 degrees of freedom gives a P-value of 0.5351. This means we can't conclude an effect of type of cartridge on the effect of distance on the spread.