Carlo Cavicchia

Recitation 2

carlo.cavicchia@uniroma1.it 🖂

An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

> Identify the response variable and the explanatory variable.

An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

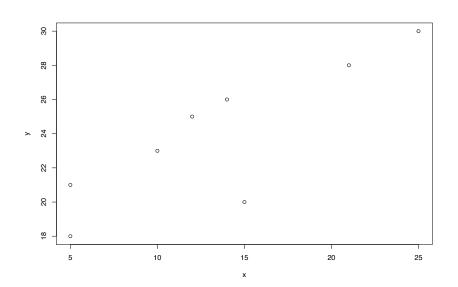
Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

Response variable :: Grade - Explanatory variable :: Study time

An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

> Construct a scatterplot.



An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

> Find and interpret the correlation coefficient.

To compute the correlation we should start from the means: $\bar{x}=13.375$, $\bar{y}=23.875$

$x-\bar{x}$	$y - \bar{y}$	$(x-\bar{x})\times(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
0.62	2.12	1.31	0.38	4.49
11.62	6.12	71.11	135.02	37.45
1.62	-3.88	-6.29	2.62	15.05
-8.38	-5.88	49.27	70.22	34.57
-3.38	-0.88	2.97	11.42	0.77
-1.38	1.12	-1.55	1.90	1.25
-8.38	-2.88	24.13	70.22	8.29
7.62	4.12	31.39	58.06	16.97
		172.34	349.84	118.84

It follows that $r=172.34/\sqrt{(349.84\times118.84)}=0.8452$. It means that Study Time and Grade are highly positive correlated. When the Study Time increases, the Grade increases too.

An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

> If all grades increase by one, does the correlation coefficient change?

An instructor of Statistics collected data from one of her classes in Spring 2016 to investigate the relationship between Study time per week (number of hours) to predict the final grade. For the 8 students in her class the data were as shown in the table.

Student	Study Time	Grade
1	14	26
2	25	30
3	15	20
4	5	18
5	10	23
6	12	25
7	5	21
8	21	28

It will not change. FOR HOME: VERIFY IT

You flip a coin three times.

> Use a tree diagram to show the possible outcome patterns. How many outcomes are in the sample space?

up to you -> https://forms.gle/Gvq5iS9M3ZGtFaWW6

You flip a coin three times.

> Use a tree diagram to show the possible outcome patterns. How many outcomes are in the sample space?

First Coin	Second Coin	Third Coin
Т	Т	T
T	T	Н
Т	Н	Т
T	Н	Н
Н	T	T
Н	T	Н
Н	Н	T
Н	Н	Н

There are $2^3 = 8$ outcomes.

You flip a coin three times.

> Using the sample space constructed in part 1, find the probability (i) to have at least 2 heads; (ii) to have at least one tail.

You flip a coin three times.

> Using the sample space constructed in part 1, find the probability (i) to have at least 2 heads

$$P_1 = \frac{\text{favorable cases}}{\text{possible cases}} = \frac{4}{8} = 0.5$$

You flip a coin three times.

> Using the sample space constructed in part 1, find the probability (ii) to have at least one tail

up to you -> https://forms.gle/hrHWSkpXXA4bvMtB8

You flip a coin three times.

> Using the sample space constructed in part 1, find the probability (ii) to have at least one tail

$$P_2 = \frac{\text{favorable cases}}{\text{possible cases}} = \frac{7}{8}$$

A die is rolled.

> List the possible outcomes in the sample space.

A die is rolled.

> List the possible outcomes in the sample space.

$$S = \{1, 2, 3, 4, 5, 6\}$$

A die is rolled.

> What is the probabilty of getting a number which is even?

A die is rolled.

> What is the probabilty of getting a number which is even?

$$P = \frac{\#\{2,4,6\}}{\#\{1,2,3,4,5,6\}} = 0.5$$

A die is rolled.

> What is the probabilty of getting a number which is greater than 4?

A die is rolled.

> What is the probabilty of getting a number which is greater than 4?

$$P = \frac{\#\{5,6\}}{\#\{1,2,3,4,5,6\}} = 0.333$$

A die is rolled.

> What is the probabilty of getting a number which is less than 3? What is its complement?

A die is rolled.

> What is the probabilty of getting a number which is less than 3? What is its complement?

up to you -> https://forms.gle/68pmVmeAN7EKeUxi9

A die is rolled.

> What is the probabilty of getting a number which is less than 3? What is its complement?

$$P = \frac{\#\{1,2\}}{\#\{1,2,3,4,5,6\}} = 0.333$$

$$Q = \frac{\#\{3, 4, 5, 6\}}{\#\{1, 2, 3, 4, 5, 6\}} = 0.666 = 1 - 0.333 = 1 - P$$

Two dice are rolled.

> Construct the sample space. How many outcomes are there?

Two dice are rolled.

> Construct the sample space. How many outcomes are there?

	1	2 3 4 5 6	3	4	5	6
1	2	3	4 5	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

There are $6^2 = 36$ possible outcomes.

Two dice are rolled.

> Find the probability of rolling a sum of 7.

Two dice are rolled.

> Find the probability of rolling a sum of 7.

$$P = \frac{\#\{(1,6),(6,1),(3,4),(4,3),(2,5),(5,2)\}}{36} = 0.1667$$

Two dice are rolled.

> Find the probability of getting a total of at least 10.

Two dice are rolled.

 $\,\,{}^{\scriptscriptstyle >}$ Find the probability of getting a total of at least 10.

up to you -> https://forms.gle/A559nEVe9REDa9K6A

Two dice are rolled.

> Find the probability of getting a total of at least 10.

$$P = \frac{\#\{(6,4),(4,6),(5,5),(6,5),(5,6),(6,6)\}}{36} = 0.1667$$

Two dice are rolled.

> Find the probability of getting a odd number as the sum.

Two dice are rolled.

> Find the probability of getting a odd number as the sum.

$$P = \frac{18}{36} = 0.5$$

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

> What is the sample space of possible outcomes for a randomly selected individual involved in an auto accident?

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

SY - SN - DY - DN

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

> Compute (i) P(D), (ii) P(N).

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Survived (S)	Died (D)
410	5
160	15
	410

$$P(D) = \frac{20}{590} = 0.03$$

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

$$P(N) = \frac{175}{590} = 0.30$$

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

> Compute the probability that an individual did not wear a seat belt and survived.

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

$$P(N \cap S) = 160/590 = 0.27$$

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

Are the events N and S independent?

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

If events N and S were independent, then

$$P(N \cap S) = P(N) \times P(S) = \frac{175}{590} \times \frac{570}{590} = 0.299 \times 0.97 = 0.291$$

So N and S are not independent. This indicates that chance of surviving depends on seat belt use since 0.291 is not equal to 0.27.

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

> Compute the probability that the individual survived, given that the person (i) wore and (ii) did not wear a seat belt. Interpret the results.

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Survived (S)	Died (D)
410	5
160	15
	410

$$P(S \mid Y) = P(S \cap Y)/P(Y) = \frac{\frac{410}{590}}{\frac{415}{590}} = 0.99$$

$$P(S \mid N) = P(S \cap N)/P(N) = \frac{\frac{160}{590}}{\frac{175}{590}} = 0.91$$

Once again, it means that the events are not independent. Wearing or not the seat belt influences the chance of surviving.

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Survived (S)	Died (D)
410	5
160	15
	410

> Are the events of dying and wearing a seat belt independent? Justify your answer.

Based on records of automobile accidents, it has been reported the counts who survived (S) and died (D) according to whether they wore a seat belt (Y=yes, N=no). The data are presented in the following contingency table.

Wore Seat Belt	Survived (S)	Died (D)
Yes (Y)	410	5
No(N)	160	15

They are not independent since neither $P(S \mid Y)$ nor $P(S \mid N)$ equals P(S). Specifically 0.99 and 0.91 are different from 0.97.

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

> Set up a contingency table that cross classifies gender by level of happiness.

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

Gender	Very Happy	Pretty Happy	Not Too Happy	Total
Male	184	235	41	460
Female	232	265	43	540

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

> Compute the probability that a married adult is very happy.

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

up to you -> https://forms.gle/LqqQUJ3N1aUsWTvr5

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

$$P(VH) = (0.46 \times 0.4) + (0.54 \times 0.43) = \frac{184 + 232}{1000} = 0.42$$

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

> Compute the probability that a married adult is very happy, (i) given that their gender is male and (ii) given that their gender is female.

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

$$P(VH \mid M) = P(VH \cap M)/P(M) = \frac{\frac{184}{1000}}{\frac{460}{1000}} = 0.4$$

$$P(VH \mid F) = P(VH \cap F)/P(F) = \frac{\frac{232}{1000}}{\frac{540}{1000}} = 0.43$$

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

> For these subjects, are the events being very happy and being a male independent?

It has been asked to 1000 people whether they were happy in their marriages. The poll reported that 46% were men, while 54% were women. Among men, 40% declared to be very happy, 51% pretty happy, while 9% not too happy. Among women, 43% declared to be very happy, 49% pretty happy, while 8% not too happy.

No, since neither $P(VH \mid M) = 0.40$ nor $P(VH \mid F) = 0.43$ are equal to P(VH) = 0.42.

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

 \rightarrow Construct the probability distribution of X and find p.

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

$$p = 1 - 0.70 - 0.20 = 0.10.$$

x_i	$P(x_i)$
80	0.70
50	0.20
20	0.10

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

 \rightarrow Find the mean of the probability distribution of X.

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

up to you -> https://forms.gle/iUnNkXJyCQe9ZWiMA

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

$$E(X) = 80 \times 0.70 + 50 \times 0.20 + 20 \times 0.10 = 68$$

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

 \rightarrow Find the variance of the probability distribution of X.

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

up to you -> https://forms.gle/uqrZZxm1bELKmYua9

A wheat farmer living in Pennsylvania finds that his annual profit is \$80 if the summer weather is typical, \$50 if the weather is unusually dry, and \$20 if there is a severe storm that destroys much of his crop. Weather bureau records indicate that the probability is 0.70 of typical weather, 0.20 of unusually dry weather, and p of a severe storm. Let X be the farmer's profit.

$\overline{x_i}$	$P(x_i)$	$x_i - E(X)$	$p(x_i)\times (x_i-E(X))^2$
80	0.70	12	100.8
50	0.20	-18	64.8
_20	0.10	-48	230.4

The variance is 396.

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

Compute the probability that

> a randomly selected vehicle speed is greater than 73

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

$$P(X>73) = P(\frac{X-50}{15} > \frac{73-50}{15}) = P(Z>1.53) =$$

$$P(Z<-1.53) = \Phi(-1.53) = 0.063$$

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

Compute the probability that

 \rightarrow a randomly selected vehicle speed is between 40 and 73

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

Compute the probability that

ightarrow a randomly selected vehicle speed is between 40 and 73

up to you -> https://forms.gle/103ZDWen5ug6YmRF8

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

$$P(40 < X < 73) = P(\frac{40 - 50}{15} < \frac{X - 50}{15} < \frac{85 - 50}{15}) =$$

$$P(-0.67 < Z < 1.53) = \Phi(1.53) - \Phi(-0.67) = 0.9370 - 0.2525 = 0.6845$$

where

$$\Phi(1.53) = 1 - \Phi(-1.53) = 1 - 0.063 = 0.9370$$

$$\Phi(-0.67) = 1 - \Phi(0.67) = 0.2525$$

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

Compute the probability that

ightarrow a randomly selected vehicle speed is less then 85

In a population the vehicle speed distribution is well approximated by a Normal curve with mean 50 and standard deviation 15.

$$P(X<85)=P(\frac{X-50}{15}<\frac{85-50}{15})=$$

$$P(Z<2.33)=1-\Phi(-2.33)=1-0.0099=0.9901$$

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

> Compute the probability of being admitted.

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

$$P(X>125) = P(\frac{X-100}{15} > \frac{125-100}{15}) = P(Z>1.67) =$$

$$P(Z<-1.67) = \Phi(-1.67) = 0.0475$$

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

Compute the probability that a randomly selected IQ score is between 120 and 145.

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

$$P(120 < X < 145) = P(\frac{120 - 100}{15} < \frac{X - 100}{15} < \frac{145 - 100}{15}) =$$

$$P(1.33 < Z < 3) = \Phi(3) - \Phi(1.33) = 0.9987 - 0.9082 = 0.0905$$

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

Compute the probability that a randomly selected IQ score is less than 125.

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

$$P(X<125)=P(\frac{X-100}{15}<\frac{125-100}{15})=P(Z<1.67)=0.9525$$

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

> Compute the probability that a randomly selected IQ score is less than 90.

At one private college, a minimum IQ score of 125 is necessary to be considered for admission. IQ scores can be well approximated with a Normal distribution with mean of 100 and standard deviation of 15.

$$P(X<90) = P(\frac{X-100}{15} < \frac{90-100}{15}) = P(Z<-0.67) = 0.2514$$

$$P(-2 < Z < -1)$$

Let the random variable \mathcal{Z} follow a standard Normal distribution.

$$P(-2 < Z < -1) = \Phi(-1) - \Phi(-2) = 0.1587 - 0.0228 = 0.1359$$

Let the random variable ${\it Z}$ follow a standard Normal distribution.

$$P(Z > 1.52) = 1 - \Phi(1.52) = 1 - 0.9357 = 0.0643$$

$$P(-2 < Z < 0.89)$$

Let the random variable ${\it Z}$ follow a standard Normal distribution.

$$P(-2 < Z < 0.89) = \Phi(0.89) - \Phi(-2) = 0.8133 - 0.0228 = 0.7905$$

Let the random variable ${\it Z}$ follow a standard Normal distribution.

$$P(0 < Z < 2.15) = \Phi(2.15) - \Phi(0) = 0.9842 - 0.5 = 0.4842$$