Stats 60 Problem Session 1

Qian Zhao

July 14, 2020

Measure of Center, Skew

Problem 1.1

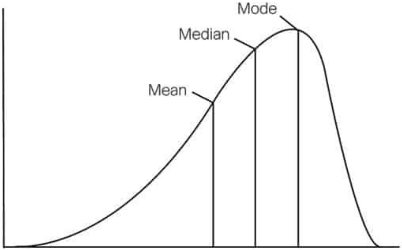
The mean undergraduate GPA at Stanford is 3.4. Do you expect (more than / less than / about) half of all undergraduates to have a GPA above 3.4 (or is it impossible to tell)?

Measure of Center, Skew

Problem 1.1

The mean undergraduate GPA at Stanford is 3.4. Do you expect (more than / less than / about) half of all undergraduates to have a GPA above 3.4 (or is it impossible to tell)?

Answer: GPA is often left skewed, so the median is right to the mean.



Left-Skewed (Negative Skewness)

Problem 2.1

What's the SD of the list $[4, 0, -2, 2, 1]$. Is it 1, 2, or 4?

Problem 2.1

What's the SD of the list $[4, 0, -2, 2, 1]$. Is it 1, 2, or 4?

Answer: It is probably 2, since the center is around 1, most data points are between distance 1 and 3 to the mean.

Measures of Spread

Problem 2.2

What is the SD of the list $[1, 3, 4, 5, 7]$?

Measures of Spread

Problem 2.2

What is the SD of the list $[1, 3, 4, 5, 7]$?

Answer:

1. The mean is $(1 + 3 + 4 + 5 + 7)/5 = 4$.

Measures of Spread

Problem 2.2

What is the SD of the list $[1, 3, 4, 5, 7]$?

Answer:

1. The mean is $(1 + 3 + 4 + 5 + 7)/5 = 4$.
2. The deviations from the mean are $[-3, -1, 0, 1, 3]$.

Measures of Spread

Problem 2.2

What is the SD of the list $[1, 3, 4, 5, 7]$?

Answer:

1. The mean is $(1 + 3 + 4 + 5 + 7)/5 = 4$.
2. The deviations from the mean are $[-3, -1, 0, 1, 3]$.
3. The mean of the squares of deviations is $(9 + 1 + 0 + 1 + 9)/5 = 4$.

Measures of Spread

Problem 2.2

What is the SD of the list $[1, 3, 4, 5, 7]$?

Answer:

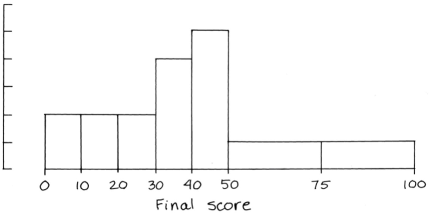
1. The mean is $(1 + 3 + 4 + 5 + 7)/5 = 4$.
2. The deviations from the mean are $[-3, -1, 0, 1, 3]$.
3. The mean of the squares of deviations is $(9 + 1 + 0 + 1 + 9)/5 = 4$.
4. The square root is 2.

Also, this is the list from problem 2.1 shifted by 3!

Histograms

Problem 3.1

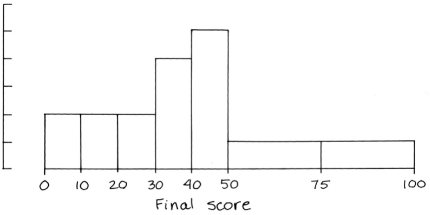
Shown below is a histogram of final exam scores. Can you estimate the 60th percentile?



Histograms

Problem 3.1

Shown below is a histogram of final exam scores. Can you estimate the 60th percentile?

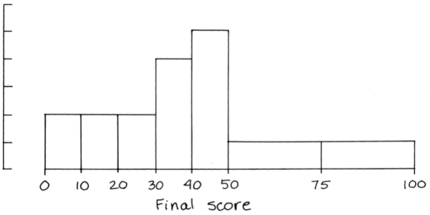


Answer: If the total area is 100%, then the area on the left of 60th percentile is 60%.

Histograms

Problem 3.1

Shown below is a histogram of final exam scores. Can you estimate the 60th percentile?



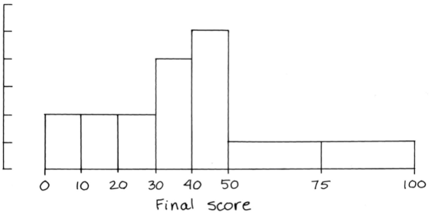
Answer: If the total area is 100%, then the area on the left of 60th percentile is 60%.

Let's say the first rectangle represents an area of a. Then the total area is $10 \times a$, which means $a = 10\%$.

Histograms

Problem 3.1

Shown below is a histogram of final exam scores. Can you estimate the 60th percentile?



Answer: If the total area is 100%, then the area on the left of 60th percentile is 60%.

Let's say the first rectangle represents an area of a. Then the total area is $10 \times a$, which means $a = 10\%$.

The 60th percentile is 44.

Normal Table

A NORMAL TABLE

z	Height	Area	z	Height	Area	z	Height	Area
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	[illegible]

Normal Curves and the Empirical Rule

Problem 4.1

IQ scores follow the normal curve with mean 100 and SD 15. People with an IQ between 115 and 130 are classified as “bright”. What percentage falls into this category?

Normal Curves and the Empirical Rule

Problem 4.1

IQ scores follow the normal curve with mean 100 and SD 15. People with an IQ between 115 and 130 are classified as "bright". What percentage falls into this category?

Answer Equivalent to area between 1 - 2 SD above the mean, so the area is

$$(95\% - 68\%)/2 = 13.5\%$$

Normal Curves and the Empirical Rule

Problem 4.2

The speed limit on the freeway is 65mph. Because of error in the radar gun readings, officers will not stop cars unless they are driving over 71mph. The police chief says that this ensures that no more than 2.5% of cars driving at the speed limit will be pulled over for speeding. Assuming radar gun readings follow a normal curve, what does this say about the SD of the readings?

Normal Curves and the Empirical Rule

Problem 4.2

The speed limit on the freeway is 65mph. Because of error in the radar gun readings, officers will not stop cars unless they are driving over 71mph. The police chief says that this ensures that no more than 2.5% of cars driving at the speed limit will be pulled over for speeding. Assuming radar gun readings follow a normal curve, what does this say about the SD of the readings?

Answer This means 71 is 2SD above the mean radar gun reading at the speed limit, which is 65, thus

$$\mathrm{SD} = (71 - 65)/2 = 3.$$

Probability Rules

Problem 5.1

Tversky and Kahneman (1982) asked subjects the following question. Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

Probability Rules

Problem 5.1

Tversky and Kahneman (1982) asked subjects the following question. Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

Answer Linda may be a bank teller without being active in the feminist movement. Therefore, it is more likely that she is a bank teller.

Probability Rules

Problem 5.1

Tversky and Kahneman (1982) asked subjects the following question. Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

Answer Linda may be a bank teller without being active in the feminist movement. Therefore, it is more likely that she is a bank teller.
You can also use conditional probability rule: The probability that Linda is a bank teller and is active in the feminist movement equals the probability that Linda is active in the feminist movement given she is a bank teller times the probability that she is a bank teller, which is less than the probability that she is a bank teller.

Probability Rules

Problem 5.2

Four draws are going to be made from the box [1] [2] [2] [3] [3]. Find the chance that [2] is drawn at least once if ...

(a) ... the draws are made with replacement.

Probability Rules

Problem 5.2

Four draws are going to be made from the box | 1 | 2 | 2 | 3 | 3 |. Find the chance that | 2 | is drawn at least once if ...
(a) ... the draws are made with replacement.

Answer If the draws are with replacement, then results from each draw are independent (results from the first draw does not affect results from the second draw). The chance that | 2 | is *not* draw in any of the four times is $(3/5)^4$. Use the complement rule to get the answer is $1 - (3/5)^4 = 0.87$.

Probability Rules

Problem 5.2

Four draws are going to be made from the box $\boxed{\boxed{1}\ \boxed{2}\ \boxed{2}\ \boxed{3}\ \boxed{3}}$. Find the chance that $\boxed{2}$ is drawn at least once if ...
(a) ... the draws are made with replacement.

Answer If the draws are with replacement, then results from each draw are independent (results from the first draw does not affect results from the second draw). The chance that $\boxed{2}$ is *not* draw in any of the four times is $(3/5)^4$. Use the complement rule to get the answer is $1-(3/5)^4 = 0.87$.

(b) ... the draws are made without replacement.

Probability Rules

Problem 5.2

Four draws are going to be made from the box [1 | 2 | 2 | 3 | 3]. Find the chance that [2] is drawn at least once if ...
(a) ... the draws are made with replacement.

Answer If the draws are with replacement, then results from each draw are independent (results from the first draw does not affect results from the second draw). The chance that [2] is *not* draw in any of the four times is $(3/5)^4$. Use the complement rule to get the answer is $1 - (3/5)^4 = 0.87$.

(b) ... the draws are made without replacement.

Answer There are only three tickets which are not a [2], so if four tickets are drawn, at least one must be a [2].

Probability Rules

Problem 5.3

10% of employees at a department store have been skimming money from the cash register. The manager decides to subject all employees to a lie detector test. The lie detector goes off 80% of the time when a person is lying, but it also goes off 25% of the time when a person is telling the truth. The lie detector beeps for a worker who claims he didn't do it. What's the chance he's lying?

Answer Let's call L be the event the detector test positive and S be the event the employee is skimming money.

$$\Pr(S \mid L)$$

Answer Let's call L be the event the detector test positive and S be the event the employee is skimming money.

$$\Pr(S \mid L) = \frac{\Pr(S \cap L)}{\Pr(L)}$$

Answer Let's call L be the event the detector test positive and S be the event the employee is skimming money.

$$\begin{aligned}\Pr(S \mid L) &= \frac{\Pr(S \cap L)}{\Pr(L)} \\ &= \frac{\Pr(L \mid S)\Pr(S)}{\Pr(L)}\end{aligned}$$

Answer Let's call L be the event the detector test positive and S be the event the employee is skimming money.

$$\begin{aligned}
\Pr(S \mid L) &= \frac{\Pr(S \cap L)}{\Pr(L)} \\
&= \frac{\Pr(L \mid S)\Pr(S)}{\Pr(L)} \\
&= \frac{\Pr(L \mid S)\Pr(S)}{\Pr(L \mid S)\Pr(S) + \Pr(L \mid \text{not } S)\Pr(\text{not} S)}
\end{aligned}$$

Answer Let's call L be the event the detector test positive and S be the event the employee is skimming money.

$$\begin{aligned}
\Pr(S \mid L) &= \frac{\Pr(S \cap L)}{\Pr(L)} \\
&= \frac{\Pr(L \mid S)\Pr(S)}{\Pr(L)} \\
&= \frac{\Pr(L \mid S)\Pr(S)}{\Pr(L \mid S)\Pr(S) + \Pr(L \mid \text{not } S)\Pr(\text{not} S)} \\
&= \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.25 \times 0.9} \\
&= 0.26.
\end{aligned}$$

Probability Rules

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Probability Rules

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Answer For any particular denomination of the first four cards, there are $4 \times 3 \times 2 \times 1 = 24$ ways to deal these cards.

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Answer For any particular denomination of the first four cards, there are $4 \times 3 \times 2 \times 1 = 24$ ways to deal these cards. Since there are 13 denominations, there are 13×24 ways in total .

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Answer For any particular denomination of the first four cards, there are $4 \times 3 \times 2 \times 1 = 24$ ways to deal these cards. Since there are 13 denominations, there are 13×24 ways in total . There are 48 ways to pick the last card.

Probability Rules

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Answer For any particular denomination of the first four cards, there are $4 \times 3 \times 2 \times 1 = 24$ ways to deal these cards. Since there are 13 denominations, there are 13×24 ways in total . There are 48 ways to pick the last card. The total number of ways to deal 5 cards is $52 \times 51 \times 50 \times 49 \times 48$.

Probability Rules

Problem 5.4

A poker hand of 5 cards is dealt from a single deck of 52 cards.
(a) What's the probability the first four cards are the same rank?

Answer For any particular denomination of the first four cards, there are $4 \times 3 \times 2 \times 1 = 24$ ways to deal these cards. Since there are 13 denominations, there are 13×24 ways in total . There are 48 ways to pick the last card. The total number of ways to deal 5 cards is $52 \times 51 \times 50 \times 49 \times 48$. Each way is equally likely, so we can use the counting principle to get

$$13 \times 24 \times 48/(52 \times 51 \times 50 \times 49 \times 48) = 5 \times 10^{-5}.$$

Probability Rules

Problem 5.4

(b) What's the probability you get "four of a kind" (four cards of the same rank)?

Probability Rules

Problem 5.4

(b) What's the probability you get "four of a kind" (four cards of the same rank)?

Answer The card of a different denomination can be placed in any of five positions.

Problem 5.4

(b) What's the probability you get "four of a kind" (four cards of the same rank)?

Answer The card of a different denomination can be placed in any of five positions. The chance that four cards of the same denomination are drawn, with the card of a different denomination in a fixed position is given by the probability calculated in the previous question.

Problem 5.4

(b) What's the probability you get "four of a kind" (four cards of the same rank)?

Answer The card of a different denomination can be placed in any of five positions. The chance that four cards of the same denomination are drawn, with the card of a different denomination in a fixed position is given by the probability calculated in the previous question. The answer is

$$5 \times 13 \times \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49} \times \frac{48}{48} = 0.0002.$$

Problem 5.5

You are in the middle of an SAT verbal section when the proctor calls out, "One minute remaining!" Oh no! You haven't even read the last passages, and there's only time to guess the answer to the 4 remaining questions. Each question has five answer.

(a) What's the chance you get all 4 questions wrong?

Problem 5.5

You are in the middle of an SAT verbal section when the proctor calls out, "One minute remaining!" Oh no! You haven't even read the last passages, and there's only time to guess the answer to the 4 remaining questions. Each question has five answer.

(a) What's the chance you get all 4 questions wrong?

Answer Whether a question is right or not are independent of each other, so we use the multiplication rule to get

$$\Pr(\text{all 4 wrong}) = \Pr(\text{question 1 wrong}) \times \ldots \times \Pr(\text{question 4 wrong})$$

Problem 5.5

You are in the middle of an SAT verbal section when the proctor calls out, "One minute remaining!" Oh no! You haven't even read the last passages, and there's only time to guess the answer to the 4 remaining questions. Each question has five answer.

(a) What's the chance you get all 4 questions wrong?

Answer Whether a question is right or not are independent of each other, so we use the multiplication rule to get

$$\begin{aligned} \Pr(\text{all 4 wrong}) &= \Pr(\text{question 1 wrong}) \times \ldots \times \Pr(\text{question 4 wrong}) \\ &= \left(\frac{4}{5}\right)^4 = 0.41. \end{aligned}$$

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$.

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability,

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability, so the chance to get one question correct is $4 \times 0.0432 = 0.1728$.

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability, so the chance to get one question correct is $4 \times 0.0432 = 0.1728$.
(c) The chance you get the questions 1 and 2 correct and questions 3 and 4 wrong is $(1/5)^2(4/5)^2 = 0.0256$.

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability, so the chance to get one question correct is $4 \times 0.0432 = 0.1728$.
(c) The chance you get the questions 1 and 2 correct and questions 3 and 4 wrong is $(1/5)^2(4/5)^2 = 0.0256$. There are 6 ways to choose 2 correct questions out of 4 questions (1&2, 1&3, 1&4, 2&3, 2&4, 3&4).

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability, so the chance to get one question correct is $4 \times 0.0432 = 0.1728$.
(c) The chance you get the questions 1 and 2 correct and questions 3 and 4 wrong is $(1/5)^2(4/5)^2 = 0.0256$. There are 6 ways to choose 2 correct questions out of 4 questions (1&2, 1&3, 1&4, 2&3, 2&4, 3&4). Thus the chance to get 2 questions correct is $6 \times 0.0256 = 0.1536$.

Probability Rules

(b) What's the probability you get exactly 1 correct?
(c) What's the probability you get exactly 2 correct?
(d) What's the probability you get any correct?

Answer (b) The chance you get the first question correct and the others wrong is $1/5 \times (4/5)^3 = 0.0432$. There are four locations this correct question could be (first, second, third or fourth), each has equal probability, so the chance to get one question correct is $4 \times 0.0432 = 0.1728$.
(c) The chance you get the questions 1 and 2 correct and questions 3 and 4 wrong is $(1/5)^2(4/5)^2 = 0.0256$. There are 6 ways to choose 2 correct questions out of 4 questions (1&2, 1&3, 1&4, 2&3, 2&4, 3&4). Thus the chance to get 2 questions correct is $6 \times 0.0256 = 0.1536$.
(d) Use the complement rule to get

$$\Pr(\text{any correct}) = 1 - \Pr(\text{all 4 wrong}) = 1 - 0.41 = 0.59.$$

Probability Rule

Problem 5.6

You and your friends want to go to a concert. Because it's very popular, each of you only have $1/3$ chance of getting the ticket if you put in an order (one person can purchase for the group).

(a) What is the probability you can successfully buy a ticket for the concert if two of you order? What about three of you order?

Probability Rule

Problem 5.6

You and your friends want to go to a concert. Because it's very popular, each of you only have $1/3$ chance of getting the ticket if you put in an order (one person can purchase for the group).
(a) What is the probability you can successfully buy a ticket for the concert if two of you order? What about three of you order?

(a) The probability the two of you *fail* to get a ticket is

$$\left(\frac{2}{3}\right)^2 = 0.44.$$

(by multiplication rule).

Probability Rule

Problem 5.6

You and your friends want to go to a concert. Because it's very popular, each of you only have $1/3$ chance of getting the ticket if you put in an order (one person can purchase for the group).
(a) What is the probability you can successfully buy a ticket for the concert if two of you order? What about three of you order?

(a) The probability the two of you *fail* to get a ticket is

$$\left(\frac{2}{3}\right)^2 = 0.44.$$

(by multiplication rule). By the complement rule, the success probability is $1 - 0.44 = 0.56$.

Probability Rule

Problem 5.6

You and your friends want to go to a concert. Because it's very popular, each of you only have 1/3 chance of getting the ticket if you put in an order (one person can purchase for the group).
(a) What is the probability you can successfully buy a ticket for the concert if two of you order? What about three of you order?

(a) The probability the two of you *fail* to get a ticket is

$$\left(\frac{2}{3}\right)^2 = 0.44.$$

(by multiplication rule). By the complement rule, the success probability is $1 - 0.44 = 0.56$.
(b) The chance that all of you fail to get a ticket is $(2/3)^3 = 0.296$. So the success probability is $1 - 0.296 = 0.704$.

Probability Rule

(c)How many of your friends plus you need to put in orders to guarantee at least 85% chance to obtain a ticket for the group?

Probability Rule

(c)How many of your friends plus you need to put in orders to guarantee at least 85% chance to obtain a ticket for the group?

Answer (c) We can calculate the success probability if 4 or 5 of you put in orders, and the success probabilities are 80% and 87% respectively. Thus if 5 of you put in an order, the success probability is at least 85%.