How big can the tail be?

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Not a lot.

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Not a lot.

But we don't have enough information to say exactly what proportion,

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Not a lot.

But we don't have enough information to say exactly what proportion, or even approximately what proportion.

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Not a lot.

But we don't have enough information to say exactly what proportion, or even approximately what proportion.

However . . .

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Not a lot.

But we don't have enough information to say exactly what proportion, or even approximately what proportion.

However . . .

Andrey Markov (1856-1922)

How big can the tail be?

The average age of a group of people is 20 years. What proportion are more than 80 years old?

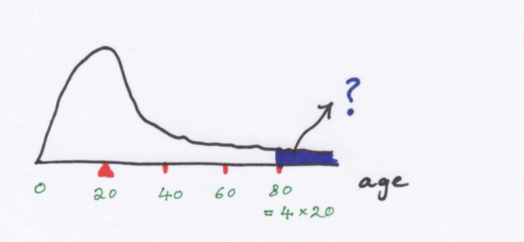
Not a lot.

But we don't have enough information to say exactly what proportion, or even approximately what proportion.

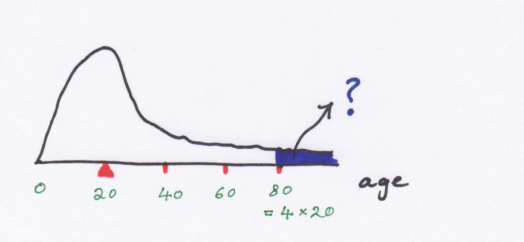
However . . .

Andrey Markov (1856-1922) came up with a simple bound.

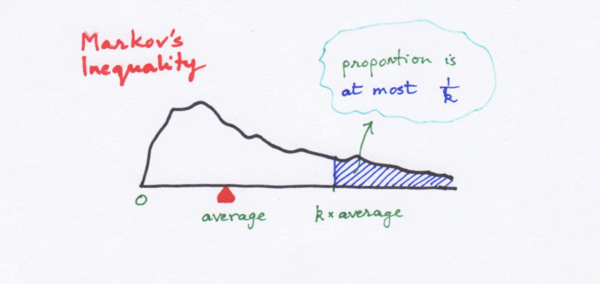
A tail bound



A tail bound



?
0
20
40
60
80
= 4 × 20
age



Markov's Inequality
proportion is at most $\frac{1}{k}$
0
average
k × average

A formal statement

A formal statement

Let k be any positive number.

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to
Math symbol: $\geq$

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to
Math symbol: $\geq$

at most means the same thing as no more than

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to
Math symbol: $\geq$

at most means the same thing as no more than

Example: In any list of non-negative numbers,

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to
Math symbol: $\geq$

at most means the same thing as no more than

Example: In any list of non-negative numbers,
the proportion of entries that are at least as large as 4 times the average

A formal statement

Let k be any positive number.

Markov's inequality

If a list has only non-negative entries, then the proportion of entries that are at least as large as k times the average is at most $1/k$.

at least as large as means the same thing as greater than or equal to
Math symbol: $\geq$

at most means the same thing as no more than

Example: In any list of non-negative numbers,
the proportion of entries that are at least as large as 4 times the average is **at most 1/4**; in other words, **no more than 25%**.

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

This is a **bound**, not an exact answer, nor an approximation, nor an estimate.

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

This is a **bound**, not an exact answer, nor an approximation, nor an estimate.

The proportion could be exactly $1/10$;

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

This is a **bound**, not an exact answer, nor an approximation, nor an estimate.

The proportion could be exactly $1/10$;
it could be quite a bit less than $1/10$;

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

This is a **bound**, not an exact answer, nor an approximation, nor an estimate.

The proportion could be exactly $1/10$;
it could be quite a bit less than $1/10$;
it could even be 0;

Understanding bounds

In any non-negative list,
the proportion of entries that are at least as large as 10 times the average can be **no more than** $1/10$.

This is a **bound**, not an exact answer, nor an approximation, nor an estimate.

The proportion could be exactly $1/10$;
it could be quite a bit less than $1/10$;
it could even be 0;
but it **cannot be more than 1/10.**

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Question: more than 80 years old: > 80

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Question: more than 80 years old: > 80
Markov: greater than or equal to 80 years old: ≥ 80

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Question: more than 80 years old: > 80
Markov: greater than or equal to 80 years old: ≥ 80

The proportion that are more than 80 years old

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Question: more than 80 years old: > 80
Markov: greater than or equal to 80 years old: ≥ 80

The proportion that are more than 80 years old is at most the proportion that are greater than or equal to 80 years old.

Taking care of the edge

The average age of a group of people is 20 years. What proportion are more than 80 years old?

Can't say exactly, or even approximately, but we can try Markov:

"80 is 4 times the average, so the proportion is **at most 1/4**"

Question: more than 80 years old: > 80
Markov: greater than or equal to 80 years old: ≥ 80

The proportion that are more than 80 years old is at most the proportion that are greater than or equal to 80 years old.

And that proportion is at most $1/4$, by Markov's inequality.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem. Markov is not wrong.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$. **????**

This is not a problem. Markov is not wrong.
Every proportion in the world is at most 200%, since proportions are at most 100%.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem. Markov is not wrong.
Every proportion in the world is at most 200%, since proportions are at most 100%.

When k is a fraction, Markov's bound is still correct as a bound.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem. Markov is not wrong.
Every proportion in the world is at most 200%, since proportions are at most 100%.

When k is a fraction, Markov's bound is still correct as a bound.
It's just not useful.

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem. Markov is not wrong.
Every proportion in the world is at most 200%, since proportions are at most 100%.

When k is a fraction, Markov's bound is still correct as a bound.
It's just not useful.

Markov's bound is most useful when k is large,

Is the bound always useful?

In a non-negative list, what proportion of entries are at least as large as half the average?

$k = 0.5$, so Markov says the proportion is at most $1/0.5 = 2 = 200\%$.
????

This is not a problem. Markov is not wrong.
Every proportion in the world is at most 200%, since proportions are at most 100%.

When k is a fraction, Markov's bound is still correct as a bound.
It's just not useful.

Markov's bound is most useful when k is large, that is, when you're interested in entries that are quite far above average.