

18.440: Lecture 32

Strong law of large numbers and Jensen's inequality

Scott Sheffield

MIT

A story about Pedro

Strong law of large numbers

Jensen's inequality

A story about Pedro

Strong law of large numbers

Jensen's inequality

Pedro's hopes and dreams

- ▶ Pedro is considering two ways to invest his life savings.

Pedro's hopes and dreams

- ▶ Pedro is considering two ways to invest his life savings.
- ▶ One possibility: put the entire sum in government insured interest-bearing savings account. He considers this completely risk free. The (post-tax) interest rate equals the inflation rate, so the real value of his savings is guaranteed not to change.

Pedro's hopes and dreams

- ▶ Pedro is considering two ways to invest his life savings.
- ▶ One possibility: put the entire sum in government insured interest-bearing savings account. He considers this completely risk free. The (post-tax) interest rate equals the inflation rate, so the real value of his savings is guaranteed not to change.
- ▶ Riskier possibility: put sum in investment where every month *real* value goes up 15 percent with probability .53 and down 15 percent with probability .47 (independently of everything else).

Pedro's hopes and dreams

- ▶ Pedro is considering two ways to invest his life savings.
- ▶ One possibility: put the entire sum in government insured interest-bearing savings account. He considers this completely risk free. The (post-tax) interest rate equals the inflation rate, so the real value of his savings is guaranteed not to change.
- ▶ Riskier possibility: put sum in investment where every month *real* value goes up 15 percent with probability .53 and down 15 percent with probability .47 (independently of everything else).
- ▶ How much does Pedro make in expectation over 10 years with risky approach? 100 years?

Pedro's hopes and dreams

- ▶ How much does Pedro make in expectation over 10 years with risky approach? 100 years?

Pedro's hopes and dreams

- ▶ How much does Pedro make in expectation over 10 years with risky approach? 100 years?
- ▶ Answer: let R_i be i.i.d. random variables each equal to 1.15 with probability .53 and .85 with probability .47. Total value after n steps is initial investment times
$$T_n := R_1 \times R_2 \times \dots \times R_n.$$

Pedro's hopes and dreams

- ▶ How much does Pedro make in expectation over 10 years with risky approach? 100 years?
- ▶ Answer: let R_i be i.i.d. random variables each equal to 1.15 with probability .53 and .85 with probability .47. Total value after n steps is initial investment times
$$T_n := R_1 \times R_2 \times \dots \times R_n.$$
- ▶ Compute $E[R_1] = .53 \times 1.15 + .47 \times .85 = 1.009$.

Pedro's hopes and dreams

- ▶ How much does Pedro make in expectation over 10 years with risky approach? 100 years?
- ▶ Answer: let R_i be i.i.d. random variables each equal to 1.15 with probability .53 and .85 with probability .47. Total value after n steps is initial investment times
$$T_n := R_1 \times R_2 \times \dots \times R_n.$$
- ▶ Compute $E[R_1] = .53 \times 1.15 + .47 \times .85 = 1.009$.
- ▶ Then $E[T_{120}] = 1.009^{120} \approx 2.93$. And $E[T_{1200}] = 1.009^{1200} \approx 46808.9$

Pedro's financial planning

- ▶ How would you advise Pedro to invest over the next 10 years if Pedro wants to be completely sure that he doesn't lose money?

Pedro's financial planning

- ▶ How would you advise Pedro to invest over the next 10 years if Pedro wants to be completely sure that he doesn't lose money?
- ▶ What if Pedro is willing to accept substantial risk if it means there is a good chance it will enable his grandchildren to retire in comfort 100 years from now?

Pedro's financial planning

- ▶ How would you advise Pedro to invest over the next 10 years if Pedro wants to be completely sure that he doesn't lose money?
- ▶ What if Pedro is willing to accept substantial risk if it means there is a good chance it will enable his grandchildren to retire in comfort 100 years from now?
- ▶ What if Pedro wants the money for himself in ten years?

Pedro's financial planning

- ▶ How would you advise Pedro to invest over the next 10 years if Pedro wants to be completely sure that he doesn't lose money?
- ▶ What if Pedro is willing to accept substantial risk if it means there is a good chance it will enable his grandchildren to retire in comfort 100 years from now?
- ▶ What if Pedro wants the money for himself in ten years?
- ▶ Let's do some simulations.

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.
- ▶ Now S_n is a sum of i.i.d. random variables.

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.
- ▶ Now S_n is a sum of i.i.d. random variables.
- ▶ $E[X_1] = E[\log R_1] = .53(\log 1.15) + .47(\log .85) \approx -.0023$.

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.
- ▶ Now S_n is a sum of i.i.d. random variables.
- ▶ $E[X_1] = E[\log R_1] = .53(\log 1.15) + .47(\log .85) \approx -.0023$.
- ▶ By the law of large numbers, if we take n extremely large, then $S_n/n \approx -.0023$ with high probability.

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.
- ▶ Now S_n is a sum of i.i.d. random variables.
- ▶ $E[X_1] = E[\log R_1] = .53(\log 1.15) + .47(\log .85) \approx -.0023$.
- ▶ By the law of large numbers, if we take n extremely large, then $S_n/n \approx -.0023$ with high probability.
- ▶ This means that, when n is large, S_n is *usually* a very negative value, which means T_n is *usually* very close to zero (even though its expectation is very large).

Logarithmic point of view

- ▶ We wrote $T_n = R_1 \times \dots \times R_n$. Taking logs, we can write $X_i = \log R_i$ and $S_n = \log T_n = \sum_{i=1}^n X_i$.
- ▶ Now S_n is a sum of i.i.d. random variables.
- ▶ $E[X_1] = E[\log R_1] = .53(\log 1.15) + .47(\log .85) \approx -.0023$.
- ▶ By the law of large numbers, if we take n extremely large, then $S_n/n \approx -.0023$ with high probability.
- ▶ This means that, when n is large, S_n is *usually* a very negative value, which means T_n is *usually* very close to zero (even though its expectation is very large).
- ▶ Bad news for Pedro's grandchildren. After 100 years, the portfolio is probably in bad shape. But what if Pedro takes an even longer view? Will T_n converge to zero with probability one as n gets large? Or will T_n perhaps always *eventually* rebound?

A story about Pedro

Strong law of large numbers

Jensen's inequality

A story about Pedro

Strong law of large numbers

Jensen's inequality

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .
- ▶ Then the value $A_n := \frac{X_1+X_2+\dots+X_n}{n}$ is called the *empirical average* of the first n trials.

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .
- ▶ Then the value $A_n := \frac{X_1 + X_2 + \dots + X_n}{n}$ is called the *empirical average* of the first n trials.
- ▶ Intuition: when n is large, A_n is typically close to μ .

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .
- ▶ Then the value $A_n := \frac{X_1 + X_2 + \dots + X_n}{n}$ is called the *empirical average* of the first n trials.
- ▶ Intuition: when n is large, A_n is typically close to μ .
- ▶ Recall: **weak law of large numbers** states that for all $\epsilon > 0$ we have $\lim_{n \rightarrow \infty} P\{|A_n - \mu| > \epsilon\} = 0$.

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .
- ▶ Then the value $A_n := \frac{X_1 + X_2 + \dots + X_n}{n}$ is called the *empirical average* of the first n trials.
- ▶ Intuition: when n is large, A_n is typically close to μ .
- ▶ Recall: **weak law of large numbers** states that for all $\epsilon > 0$ we have $\lim_{n \rightarrow \infty} P\{|A_n - \mu| > \epsilon\} = 0$.
- ▶ The **strong law of large numbers** states that with probability one $\lim_{n \rightarrow \infty} A_n = \mu$.

Strong law of large numbers

- ▶ Suppose X_i are i.i.d. random variables with mean μ .
- ▶ Then the value $A_n := \frac{X_1 + X_2 + \dots + X_n}{n}$ is called the *empirical average* of the first n trials.
- ▶ Intuition: when n is large, A_n is typically close to μ .
- ▶ Recall: **weak law of large numbers** states that for all $\epsilon > 0$ we have $\lim_{n \rightarrow \infty} P\{|A_n - \mu| > \epsilon\} = 0$.
- ▶ The **strong law of large numbers** states that with probability one $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ It is called “strong” because it implies the weak law of large numbers. But it takes a bit of thought to see why this is the case.

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ Strong law implies that for every ϵ the random variable $Y_\epsilon = \max\{n : |A_n - \mu| > \epsilon\}$ is finite with probability one. It has some probability mass function (though we don't know what it is).

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ Strong law implies that for every ϵ the random variable $Y_\epsilon = \max\{n : |A_n - \mu| > \epsilon\}$ is finite with probability one. It has some probability mass function (though we don't know what it is).
- ▶ Note that if $|A_n - \mu| > \epsilon$ for some n value then $Y_\epsilon \geq n$.

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ Strong law implies that for every ϵ the random variable $Y_\epsilon = \max\{n : |A_n - \mu| > \epsilon\}$ is finite with probability one. It has some probability mass function (though we don't know what it is).
- ▶ Note that if $|A_n - \mu| > \epsilon$ for some n value then $Y_\epsilon \geq n$.
- ▶ Thus for each n we have $P\{|A_n - \mu| > \epsilon\} \leq P\{Y_\epsilon \geq n\}$.

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ Strong law implies that for every ϵ the random variable $Y_\epsilon = \max\{n : |A_n - \mu| > \epsilon\}$ is finite with probability one. It has some probability mass function (though we don't know what it is).
- ▶ Note that if $|A_n - \mu| > \epsilon$ for some n value then $Y_\epsilon \geq n$.
- ▶ Thus for each n we have $P\{|A_n - \mu| > \epsilon\} \leq P\{Y_\epsilon \geq n\}$.
- ▶ So $\lim_{n \rightarrow \infty} P\{|A_n - \mu| > \epsilon\} \leq \lim_{n \rightarrow \infty} P\{Y_\epsilon \geq n\} = 0$.

Strong law implies weak law

- ▶ Suppose we know that the strong law holds, i.e., with probability 1 we have $\lim_{n \rightarrow \infty} A_n = \mu$.
- ▶ Strong law implies that for every ϵ the random variable $Y_\epsilon = \max\{n : |A_n - \mu| > \epsilon\}$ is finite with probability one. It has some probability mass function (though we don't know what it is).
- ▶ Note that if $|A_n - \mu| > \epsilon$ for some n value then $Y_\epsilon \geq n$.
- ▶ Thus for each n we have $P\{|A_n - \mu| > \epsilon\} \leq P\{Y_\epsilon \geq n\}$.
- ▶ So $\lim_{n \rightarrow \infty} P\{|A_n - \mu| > \epsilon\} \leq \lim_{n \rightarrow \infty} P\{Y_\epsilon \geq n\} = 0$.
- ▶ If the right limit is zero for each ϵ (strong law) then the left limit is zero for each ϵ (weak law).

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.
- ▶ Key to proof is to bound fourth moments of A_n .

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.
- ▶ Key to proof is to bound fourth moments of A_n .
- ▶ $E[A_n^4] = n^{-4} E[S_n^4] = n^{-4} E[(X_1 + X_2 + \dots + X_n)^4]$.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.
- ▶ Key to proof is to bound fourth moments of A_n .
- ▶ $E[A_n^4] = n^{-4} E[S_n^4] = n^{-4} E[(X_1 + X_2 + \dots + X_n)^4]$.
- ▶ Expand $(X_1 + \dots + X_n)^4$. Five kinds of terms: $X_i X_j X_k X_l$ and $X_i X_j X_k^2$ and $X_i X_j^3$ and $X_i^2 X_j^2$ and X_i^4 .

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.
- ▶ Key to proof is to bound fourth moments of A_n .
- ▶ $E[A_n^4] = n^{-4} E[S_n^4] = n^{-4} E[(X_1 + X_2 + \dots + X_n)^4]$.
- ▶ Expand $(X_1 + \dots + X_n)^4$. Five kinds of terms: $X_i X_j X_k X_l$ and $X_i X_j X_k^2$ and $X_i X_j^3$ and $X_i^2 X_j^2$ and X_i^4 .
- ▶ The first three terms all have expectation zero. There are $\binom{n}{2}$ of the fourth type and n of the last type, each equal to at most K . So $E[A_n^4] \leq n^{-4} \left(6 \binom{n}{2} + n \right) K$.

Proof of strong law assuming $E[X^4] < \infty$

- ▶ Assume $K := E[X^4] < \infty$. Not necessary, but simplifies proof.
- ▶ Note: $\text{Var}[X^2] = E[X^4] - E[X^2]^2 > 0$, so $E[X^2]^2 \leq K$.
- ▶ The strong law holds for i.i.d. copies of X if and only if it holds for i.i.d. copies of $X - \mu$ where μ is a constant.
- ▶ So we may as well assume $E[X] = 0$.
- ▶ Key to proof is to bound fourth moments of A_n .
- ▶ $E[A_n^4] = n^{-4} E[S_n^4] = n^{-4} E[(X_1 + X_2 + \dots + X_n)^4]$.
- ▶ Expand $(X_1 + \dots + X_n)^4$. Five kinds of terms: $X_i X_j X_k X_l$ and $X_i X_j X_k^2$ and $X_i X_j^3$ and $X_i^2 X_j^2$ and X_i^4 .
- ▶ The first three terms all have expectation zero. There are $\binom{n}{2}$ of the fourth type and n of the last type, each equal to at most K . So $E[A_n^4] \leq n^{-4} \left(6 \binom{n}{2} + n \right) K$.
- ▶ Thus $E[\sum_{n=1}^{\infty} A_n^4] = \sum_{n=1}^{\infty} E[A_n^4] < \infty$. So $\sum_{n=1}^{\infty} A_n^4 < \infty$ (and hence $A_n \rightarrow 0$) with probability 1.

A story about Pedro

Strong law of large numbers

Jensen's inequality

A story about Pedro

Strong law of large numbers

Jensen's inequality

Jensen's inequality statement

- ▶ Let X be random variable with finite mean $E[X] = \mu$.

Jensen's inequality statement

- ▶ Let X be random variable with finite mean $E[X] = \mu$.
- ▶ Let g be a **convex** function. This means that if you draw a straight line connecting two points on the graph of g , then the graph of g lies below that line. If g is twice differentiable, then convexity is equivalent to the statement that $g''(x) \geq 0$ for all x . For a concrete example, take $g(x) = x^2$.

Jensen's inequality statement

- ▶ Let X be random variable with finite mean $E[X] = \mu$.
- ▶ Let g be a **convex** function. This means that if you draw a straight line connecting two points on the graph of g , then the graph of g lies below that line. If g is twice differentiable, then convexity is equivalent to the statement that $g''(x) \geq 0$ for all x . For a concrete example, take $g(x) = x^2$.
- ▶ **Jensen's inequality:** $E[g(X)] \geq g(E[X])$.

Jensen's inequality statement

- ▶ Let X be random variable with finite mean $E[X] = \mu$.
- ▶ Let g be a **convex** function. This means that if you draw a straight line connecting two points on the graph of g , then the graph of g lies below that line. If g is twice differentiable, then convexity is equivalent to the statement that $g''(x) \geq 0$ for all x . For a concrete example, take $g(x) = x^2$.
- ▶ **Jensen's inequality:** $E[g(X)] \geq g(E[X])$.
- ▶ Similarly, if g is **concave** (which means $-g$ is convex), then $E[g(X)] \leq g(E[X])$.

Jensen's inequality statement

- ▶ Let X be random variable with finite mean $E[X] = \mu$.
- ▶ Let g be a **convex** function. This means that if you draw a straight line connecting two points on the graph of g , then the graph of g lies below that line. If g is twice differentiable, then convexity is equivalent to the statement that $g''(x) \geq 0$ for all x . For a concrete example, take $g(x) = x^2$.
- ▶ **Jensen's inequality:** $E[g(X)] \geq g(E[X])$.
- ▶ Similarly, if g is **concave** (which means $-g$ is convex), then $E[g(X)] \leq g(E[X])$.
- ▶ If your utility function is concave, then you always prefer a safe investment over a risky investment with the same expected return.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.
- ▶ Precisely: if X is end-of-year portfolio value, Pedro gets

$$g(X) = .02C + .2 \max\{X - C, 0\}.$$

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.
- ▶ Precisely: if X is end-of-year portfolio value, Pedro gets

$$g(X) = .02C + .2 \max\{X - C, 0\}.$$

- ▶ Pedro notices that g is a convex function. He can therefore increase his expected return by adopting risky strategies.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.
- ▶ Precisely: if X is end-of-year portfolio value, Pedro gets

$$g(X) = .02C + .2 \max\{X - C, 0\}.$$

- ▶ Pedro notices that g is a convex function. He can therefore increase his expected return by adopting risky strategies.
- ▶ Pedro has strategy that increases portfolio value 10 percent with probability .9, loses everything with probability .1.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.
- ▶ Precisely: if X is end-of-year portfolio value, Pedro gets

$$g(X) = .02C + .2 \max\{X - C, 0\}.$$

- ▶ Pedro notices that g is a convex function. He can therefore increase his expected return by adopting risky strategies.
- ▶ Pedro has strategy that increases portfolio value 10 percent with probability .9, loses everything with probability .1.
- ▶ He repeats this yearly until fund collapses.

More about Pedro

- ▶ Disappointed by the strong law of large numbers, Pedro seeks a better way to make money.
- ▶ Signs up for job as “hedge fund manager”. Allows him to manage $C \approx 10^9$ dollars of somebody else’s money. At end of each year, he and his staff get two percent of principle plus twenty percent of profit.
- ▶ Precisely: if X is end-of-year portfolio value, Pedro gets

$$g(X) = .02C + .2 \max\{X - C, 0\}.$$

- ▶ Pedro notices that g is a convex function. He can therefore increase his expected return by adopting risky strategies.
- ▶ Pedro has strategy that increases portfolio value 10 percent with probability .9, loses everything with probability .1.
- ▶ He repeats this yearly until fund collapses.
- ▶ With high probability Pedro is rich by then.

- ▶ The “two percent of principle plus twenty percent of profit” is common in the hedge fund industry.

- ▶ The “two percent of principle plus twenty percent of profit” is common in the hedge fund industry.
- ▶ The idea is that fund managers have both guaranteed revenue for expenses (two percent of principle) and incentive to make money (twenty percent of profit).

- ▶ The “two percent of principle plus twenty percent of profit” is common in the hedge fund industry.
- ▶ The idea is that fund managers have both guaranteed revenue for expenses (two percent of principle) and incentive to make money (twenty percent of profit).
- ▶ Because of Jensen’s inequality, the convexity of the payoff function is a genuine concern for hedge fund investors. People worry that it encourages fund managers (like Pedro) to take risks that are bad for the client.

- ▶ The “two percent of principle plus twenty percent of profit” is common in the hedge fund industry.
- ▶ The idea is that fund managers have both guaranteed revenue for expenses (two percent of principle) and incentive to make money (twenty percent of profit).
- ▶ Because of Jensen’s inequality, the convexity of the payoff function is a genuine concern for hedge fund investors. People worry that it encourages fund managers (like Pedro) to take risks that are bad for the client.
- ▶ This is a special case of the “principal-agent” problem of economics. How do you ensure that the people you hire genuinely share your interests?