We'll get to the CLT in a few slides, but first some preliminaries.

The tool we need to prove the CLT is the sequence of **moments** of a probability distribution.

We'll get to the CLT in a few slides, but first some preliminaries.

The tool we need to prove the CLT is the sequence of **moments** of a probability distribution.

Moments provide a way of describing distributions different from density functions or c.d.f.s, and better suited to the CLT.

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.
But if it *does* exit for *all* n, those numbers together tell you pretty much everything you need to know about X.

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.
But if it *does* exit for *all* n, those numbers together tell you pretty much everything you need to know about X.
For example,

$$\mu_X = M_1$$

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.
But if it *does* exit for *all* n, those numbers together tell you pretty much everything you need to know about X.
For example,

$$\mu_X = M_1$$

$$\sigma_X = \sqrt{M_2 - M_1^2}.$$

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.
But if it *does* exit for *all* n, those numbers together tell you pretty much everything you need to know about X.
For example,

$$\mu_X = M_1$$

$$\sigma_X = \sqrt{M_2 - M_1^2}.$$

If $g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$, then

$$E(g(X)) = a_0 + a_1 M_1 + a_2 M_2 + \cdots + a_k M_k.$$

The n^{th} moment M_n of a RV X is defined to be $\mathrm{E}(X^n)$,

$$M_n = \mathrm{E}(X^n).$$

This may not exist, especially when n is large.
But if it *does* exit for *all* n, those numbers together tell you pretty much everything you need to know about X.
For example,

$$\mu_X = M_1$$

$$\sigma_X = \sqrt{M_2 - M_1^2}.$$

If $g(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$, then

$$E(g(X)) = a_0 + a_1 M_1 + a_2 M_2 + \cdots + a_k M_k.$$

In other words we can compute the expected value of any polynomial function of X just using the M_n's.

We can even compute things like

$$\mathrm{P}(a \leq X \leq a + \mathrm{d}x).$$

How?

We can even compute things like

$$\mathrm{P}(a \leq X \leq a + \mathrm{d}x).$$

How?
Define

$$b(x) = \frac{1}{\sqrt{2\pi\epsilon}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-a}{\epsilon}\right)^2}.$$

This is just a very narrow bell curve, in which 99% of the bell is within 3ϵ of a.

We can even compute things like

$$\mathrm{P}(a \leq X \leq a + \mathrm{d}x).$$

How?
Define

$$b(x) = \frac{1}{\sqrt{2\pi\epsilon}} \mathrm{e}^{-\frac{1}{2}(\frac{x-a}{\epsilon})^2}.$$

This is just a very narrow bell curve, in which 99% of the bell is within 3ϵ of a.

Note that b is equivalent to a normal p.d.f., but I'm *not* using it here as a p.d.f., I'm using it like the g on the previous slide: I'm going to find $\mathrm{E}(b(X))$.

We can even compute things like

$$\mathrm{P}(a \leq X \leq a + \mathrm{d}x).$$

How?
Define

$$b(x) = \frac{1}{\sqrt{2\pi\epsilon}} \mathrm{e}^{-\frac{1}{2}(\frac{x-a}{\epsilon})^2}.$$

This is just a very narrow bell curve, in which 99% of the bell is within 3ϵ of a.

Note that b is equivalent to a normal p.d.f., but I'm *not* using it here as a p.d.f., I'm using it like the g on the previous slide: I'm going to find $\mathrm{E}(b(X))$.

Since b is essentially a polynomial (an infinite one, but still...) it has the same property as g: that its expectation is determined by the M_n's.

We can even compute things like

$$\mathrm{P}(a \leq X \leq a + \mathrm{d}x).$$

How?
Define

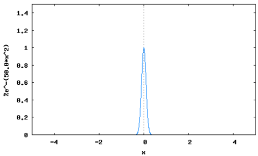
$$b(x) = \frac{1}{\sqrt{2\pi\epsilon}} \mathrm{e}^{-\frac{1}{2}(\frac{x-a}{\epsilon})^2}.$$

This is just a very narrow bell curve, in which 99% of the bell is within 3ϵ of a.

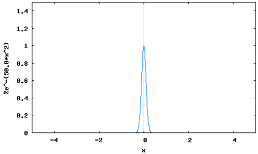
Note that b is equivalent to a normal p.d.f., but I'm *not* using it here as a p.d.f., I'm using it like the g on the previous slide: I'm going to find $\mathrm{E}(b(X))$.

Since b is essentially a polynomial (an infinite one, but still...) it has the same property as g: that its expectation is determined by the M_n's.

And the shape of b is useful because it focuses attention on a small range of x values.

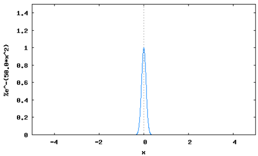


Now suppose that X is a continuous RV with p.d.f. f.
If ϵ is small enough, then we can assume that f is constant over almost the entire bell.



Now suppose that X is a continuous RV with p.d.f. f.
If ϵ is small enough, then we can assume that f is constant over almost the entire bell.
That means

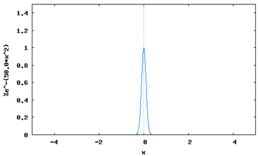
$$\mathrm{E}(b(X)) = \int_{-\infty}^{\infty} b(x) f(x) \, \mathrm{d}x$$



Now suppose that X is a continuous RV with p.d.f. f.
If ϵ is small enough, then we can assume that f is constant over almost the entire bell.
That means

$$\mathrm{E}(b(X)) = \int_{-\infty}^{\infty} b(x) f(x)\,\mathrm{d}x$$

$$= \int_{|X-a|>3\epsilon} b(x) f(x)\,\mathrm{d}x + \int_{|X-a|\leq 3\epsilon} b(x) f(x)\,\mathrm{d}x$$



Now suppose that X is a continuous RV with p.d.f. f.
If ϵ is small enough, then we can assume that f is constant over almost the entire bell.
That means

$$\mathrm{E}(b(X)) = \int_{-\infty}^{\infty} b(x) f(x)\,\mathrm{d}x$$

$$= \int_{|X-a|>3\epsilon} b(x) f(x)\,\mathrm{d}x + \int_{|X-a|\leq 3\epsilon} b(x) f(x)\,\mathrm{d}x$$

$$\approx 0 + f(a) \int_{|X-a|\leq 3\epsilon} b(x)\,\mathrm{d}x \approx f(a).$$

This means that the density function f of X is determined by the expected values of RV's like $b(X)$.

This means that the density function f of X is determined by the expected values of RV's like $b(X)$.
And that in turn means that f is determined by the M_n's, because although $b(x)$ is not a polynomial, it is a *limit* of polynomials:

$$b(x) = \frac{1}{\sqrt{2\pi\epsilon}} \mathrm{e}^{-\frac{1}{2}(\frac{x-a}{\epsilon})^2}$$

$$= \frac{1}{\sqrt{2\pi\epsilon}} \sum_{k=0}^{\infty} \frac{(-\frac{1}{2}(\frac{x-a}{\epsilon})^2)^k}{k!}$$

In a similar way it can be shown that those functions b also determine the probabilities in discrete distributions.

In a similar way it can be shown that those functions b also determine the probabilities in discrete distributions.

I'll consider it established now that the M_n's determine all the properties of a distribution. Hence if we can show that two distributions have finite and equal M_n for all n we will know they are actually the same distribution.
Now let's use the moments M_n to prove the Central Limit Theorem.

Central Limit Theorem

Theorem

Let X_i for $i = 1, 2, 3, \ldots$ be i.i.d. RVs with mean 0 and s.d. 1. Define

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i.$$

Then as $n \to \infty$ the distribution of A_n approaches the standard normal.

Central Limit Theorem

Theorem

Let X_i *for* $i = 1, 2, 3, \ldots$ *be i.i.d. RVs with mean 0 and s.d. 1. Define*

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i.$$

Then as $n \to \infty$ *the distribution of* A_n *approaches the standard normal.*

Note that the particular distribution of the X_i's *does not matter* — the limiting distribution is the same no matter what you start out with.

The basic idea is to show that the moments of the limiting distribution depend only on the first and second moments of the X_i — nothing else.

The basic idea is to show that the moments of the limiting distribution depend only on the first and second moments of the X_i — nothing else.

Let's warm up by computing a few easy moments of A_n.

$$\mathrm{E}(A_n) = \mathrm{E}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right) = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}\mathrm{E}(X_i)\right) = 0$$

because all the X_i have mean 0.

Second moment:

$$\mathrm{E}(A_n^2) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^2\right) = \frac{1}{n}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^2) + \sum_{i\neq j}\mathrm{E}(X_i X_j)\right]$$

Second moment:

$$\mathrm{E}(A_n^2) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^2\right) = \frac{1}{n}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^2) + \sum_{i\neq j}\mathrm{E}(X_i X_j)\right]$$

$$= \frac{1}{n}[n+0] = 1.$$

Second moment:

$$\mathrm{E}(A_n^2) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^2\right) = \frac{1}{n}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^2) + \sum_{i\neq j}\mathrm{E}(X_iX_j)\right]$$

$$= \frac{1}{n}[n+0] = 1.$$

So now we know that A_n has mean 0 and s.d. 1 for all n.

Second moment:

$$\mathrm{E}(A_n^2) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^2\right) = \frac{1}{n}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^2) + \sum_{i \neq j}\mathrm{E}(X_i X_j)\right]$$

$$= \frac{1}{n}[n + 0] = 1.$$

So now we know that A_n has mean 0 and s.d. 1 for all n.

We conclude therefore the limiting distribution of the A_n also has mean 0 and s.d. 1.

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= \frac{1}{n^{3/2}}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^3) + \sum_{i\neq j}\mathrm{E}(X_i X_j^2) + \sum_{\text{distinct } i,j,k}\mathrm{E}(X_i X_j X_k)\right]$$

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= \frac{1}{n^{3/2}}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^3) + \sum_{i\neq j}\mathrm{E}(X_i X_j^2) + \sum_{\text{distinct } i,j,k}\mathrm{E}(X_i X_j X_k)\right]$$

$$= \frac{1}{n^{3/2}}\left[n\mathrm{E}(X_1^3) + n(n-1)\mathrm{E}(X_1)\mathrm{E}(X_2^2) + \binom{n}{3}\mathrm{E}(X_1)^3\right]$$

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= \frac{1}{n^{3/2}}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^3) + \sum_{i\neq j}\mathrm{E}(X_i X_j^2) + \sum_{\text{distinct } i,j,k}\mathrm{E}(X_i X_j X_k)\right]$$

$$= \frac{1}{n^{3/2}}\left[n\mathrm{E}(X_1^3) + n(n-1)\mathrm{E}(X_1)\mathrm{E}(X_2^2) + \binom{n}{3}\mathrm{E}(X_1)^3\right]$$

$$= \frac{1}{n^{3/2}}(n\, M_3)$$

where M_3 is the third moment of X_1.

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= \frac{1}{n^{3/2}}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^3) + \sum_{i\neq j}\mathrm{E}(X_i X_j^2) + \sum_{\text{distinct } i,j,k}\mathrm{E}(X_i X_j X_k)\right]$$

$$= \frac{1}{n^{3/2}}\left[n\mathrm{E}(X_1^3) + n(n-1)\mathrm{E}(X_1)\mathrm{E}(X_2^2) + \binom{n}{3}\mathrm{E}(X_1)^3\right]$$

$$= \frac{1}{n^{3/2}}(n\,M_3)$$

where M_3 is the third moment of X_1.
But this expression goes to 0 as $n \to \infty$.

Third moment.

$$\mathrm{E}(A_n^3) = \mathrm{E}\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^3\right)$$

$$= \frac{1}{n^{3/2}}\left[\sum_{i=1}^{n}\mathrm{E}(X_i^3) + \sum_{i\neq j}\mathrm{E}(X_i X_j^2) + \sum_{\text{distinct } i,j,k}\mathrm{E}(X_i X_j X_k)\right]$$

$$= \frac{1}{n^{3/2}}\left[n\mathrm{E}(X_1^3) + n(n-1)\mathrm{E}(X_1)\mathrm{E}(X_2^2) + \binom{n}{3}\mathrm{E}(X_1)^3\right]$$

$$= \frac{1}{n^{3/2}}(n\, M_3)$$

where M_3 is the third moment of X_1.
But this expression goes to 0 as $n \to \infty$.
So the third moment of the limiting distribution of A_n is 0.

Note that so far the only higher moment of the X_i that has appeared has been M_3 and that ended up cancelling as $n \longrightarrow \infty$.

Note that so far the only higher moment of the X_i that has appeared has been M_3 and that ended up cancelling as $n \longrightarrow \infty$.

One obvious pattern you'll have noticed is that any term involving a factor like $\mathrm{E}(X_i)$ drops out because that value is 0.

Note that so far the only higher moment of the X_i that has appeared has been M_3 and that ended up cancelling as $n \longrightarrow \infty$.

One obvious pattern you'll have noticed is that any term involving a factor like $\mathrm{E}(X_i)$ drops out because that value is 0.

So the only terms to worry about are those in which each factor is a second or higher moment of X_i.

Note that so far the only higher moment of the X_i that has appeared has been M_3 and that ended up cancelling as $n \longrightarrow \infty$.

One obvious pattern you'll have noticed is that any term involving a factor like $\mathrm{E}(X_i)$ drops out because that value is 0.

So the only terms to worry about are those in which each factor is a second or higher moment of X_i.

There's another pattern lurking here though...

Consider an example: Suppose we're computing the fifth moment of A_n. One expression we'll have to deal with is

$$\sum_{i \neq j} \mathrm{E}(X_i^2)\mathrm{E}(X_j^3).$$

Consider an example: Suppose we're computing the fifth moment of A_n. One expression we'll have to deal with is

$$\sum_{i \neq j} \mathrm{E}(X_i^2)\mathrm{E}(X_j^3).$$

Each summand is equal to $M_2 M_3$, and the number of summands is no more (actually fewer than) n^2.

$$\leq n^2 M_2 M_3$$

Consider an example: Suppose we're computing the fifth moment of A_n. One expression we'll have to deal with is

$$\sum_{i \neq j} \mathrm{E}(X_i^2)\mathrm{E}(X_j^3).$$

Each summand is equal to $M_2 M_3$, and the number of summands is no more (actually fewer than) n^2.

$$\leq n^2 M_2 M_3$$

Now recall that A_n includes a factor of $\frac{1}{\sqrt{n}}$, so that the fifth moment of A_n in includes as factor of $\frac{1}{\sqrt{n}^5} = \frac{1}{n^{5/2}}$.

Consider an example: Suppose we're computing the fifth moment of A_n. One expression we'll have to deal with is

$$\sum_{i \neq j} \mathrm{E}(X_i^2)\mathrm{E}(X_j^3).$$

Each summand is equal to $M_2 M_3$, and the number of summands is no more (actually fewer than) n^2.

$$\leq n^2 M_2 M_3$$

Now recall that A_n includes a factor of $\frac{1}{\sqrt{n}}$, so that the fifth moment of A_n in includes as factor of $\frac{1}{\sqrt{n}^5} = \frac{1}{n^{5/2}}$.

Multiplying these two expressions together gives $\frac{M_2 M_3}{\sqrt{n}}$ which once again approaches 0 as $n \to \infty$.

Evidently very few terms survive the limit. In fact the only expressions that don't cancel in the limit are those involving only second moments of X_i:

$$\frac{1}{n^{k/2}} \sum_{\text{distinct } i_1, i_2, \ldots, i_{k/2}} \mathrm{E}(X_{i_1}^2)\mathrm{E}(X_{i_2}^2) \ldots \mathrm{E}(X_{i_{k/2}}^2)$$

Evidently very few terms survive the limit. In fact the only expressions that don't cancel in the limit are those involving only second moments of X_i:

$$\frac{1}{n^{k/2}} \sum_{\text{distinct } i_1, i_2, \ldots, i_{k/2}} \mathrm{E}(X_{i_1}^2)\mathrm{E}(X_{i_2}^2) \ldots \mathrm{E}(X_{i_{k/2}}^2)$$

$$= \frac{1}{n^{k/2}} \binom{n}{k/2} M_2^{k/2}.$$

Evidently very few terms survive the limit. In fact the only expressions that don't cancel in the limit are those involving only second moments of X_i:

$$\frac{1}{n^{k/2}} \sum_{\text{distinct } i_1, i_2, \ldots, i_{k/2}} \mathrm{E}(X_{i_1}^2)\mathrm{E}(X_{i_2}^2)\ldots\mathrm{E}(X_{i_{k/2}}^2)$$

$$= \frac{1}{n^{k/2}} \binom{n}{k/2} M_2^{k/2}.$$

Since $\binom{n}{k/2}$ is on the order of $n^{k/2}$, this term does *not* cancel as $n \to \infty$.

Evidently very few terms survive the limit. In fact the only expressions that don't cancel in the limit are those involving only second moments of X_i:

$$\frac{1}{n^{k/2}} \sum_{\text{distinct } i_1, i_2, \ldots, i_{k/2}} \mathrm{E}(X_{i_1}^2)\mathrm{E}(X_{i_2}^2) \ldots \mathrm{E}(X_{i_{k/2}}^2)$$

$$= \frac{1}{n^{k/2}} \binom{n}{k/2} M_2^{k/2}.$$

Since $\binom{n}{k/2}$ is on the order of $n^{k/2}$, this term does *not* cancel as $n \to \infty$.

Note that even this is only possible if k is even,

Evidently very few terms survive the limit. In fact the only expressions that don't cancel in the limit are those involving only second moments of X_i:

$$\frac{1}{n^{k/2}} \sum_{\text{distinct } i_1, i_2, \ldots, i_{k/2}} \mathrm{E}(X_{i_1}^2)\mathrm{E}(X_{i_2}^2) \ldots \mathrm{E}(X_{i_{k/2}}^2)$$

$$= \frac{1}{n^{k/2}} \binom{n}{k/2} M_2^{k/2}.$$

Since $\binom{n}{k/2}$ is on the order of $n^{k/2}$, this term does *not* cancel as $n \to \infty$.

Note that even this is only possible if k is even,
so we conclude that odd moments of A_n approach 0 in the limit.

OK, that was the hard part, the rest is easy.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

Well it just so happens that we know a distribution with mean 0 and s.d. 1 which also just so happens to play very well with linear combinations of itself.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

Well it just so happens that we know a distribution with mean 0 and s.d. 1 which also just so happens to play very well with linear combinations of itself.

I refer of course to the standard normal distribution.
If the X_i are standard normal, then A_n, being a linear combination of independent standard normals is also normal.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

Well it just so happens that we know a distribution with mean 0 and s.d. 1 which also just so happens to play very well with linear combinations of itself.

I refer of course to the standard normal distribution.
If the X_i are standard normal, then A_n, being a linear combination of independent standard normals is also normal.

Moreover we calculated the mean and s.d. of A_n earlier and they turned out to be 0 and 1 respectively.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

Well it just so happens that we know a distribution with mean 0 and s.d. 1 which also just so happens to play very well with linear combinations of itself.

I refer of course to the standard normal distribution.
If the X_i are standard normal, then A_n, being a linear combination of independent standard normals is also normal.

Moreover we calculated the mean and s.d. of A_n earlier and they turned out to be 0 and 1 respectively.

I.e., if the X_i's are standard normal then so are the A_n's.

OK, that was the hard part, the rest is easy.

Now we know that the limiting distribution of A_n depends only on M_1 and M_2. That means that any X_i's with the same mean (0) and the same s.d. (1) will have the same limiting distribution.

Well it just so happens that we know a distribution with mean 0 and s.d. 1 which also just so happens to play very well with linear combinations of itself.

I refer of course to the standard normal distribution.
If the X_i are standard normal, then A_n, being a linear combination of independent standard normals is also normal.

Moreover we calculated the mean and s.d. of A_n earlier and they turned out to be 0 and 1 respectively.

I.e., if the X_i's are standard normal then so are the A_n's.

And if every single A_n is standard normal, then the limiting distribution is also.

To put this together:
Since there is *some* distribution for the X_i which produces the standard normal as the limit of A_n,

then in fact *every* distribution, once it's been standardized to have mean 0 and s.d. 1, must produce the standard normal in the limit also.

To put this together:
Since there is *some* distribution for the X_i which produces the standard normal as the limit of A_n,

then in fact *every* distribution, once it's been standardized to have mean 0 and s.d. 1, must produce the standard normal in the limit also.

That's because we have shown that the limiting distribution depends only on the mean and s.d. of the original distribution.

To put this together:
Since there is *some* distribution for the X_i which produces the standard normal as the limit of A_n,

then in fact *every* distribution, once it's been standardized to have mean 0 and s.d. 1, must produce the standard normal in the limit also.

That's because we have shown that the limiting distribution depends only on the mean and s.d. of the original distribution.

Q.E.D.