Central limit theorem

Let $X_1, X_2, \ldots, X_n$ be independent + identically distubuted (i.i.d.) with $E[X_k] = \mu_X$, $Var(X_k) = \sigma_X^2$.

Define $W = X_1 + X_2 + \cdots + X_n$.

Then ~~that~~ $E[W] = n\mu_X$, $Var(W) = n\sigma_X^2$.

$$f_W(w) = f_{X_1}(w) * f_{X_2}(w) * \cdots * f_{X_n}(w)$$

$\hookrightarrow$ As n increases, this pdf $f_W(w)$ starts to look Gaussian!

LOOK AT PICS.

EX: A modem transmits one million bits. Each bit is 0 or 1 independently with equal probability. Estimate the probability of seeing at least 502,000 ones.

Sol: This is a binomial RV with $n = 1{,}000{,}000$. We would need to sum hundreds of thousands of terms! (to get this exactly)

TRICK

Let X_i = value of bit i = $\begin{cases} 0 & \text{prob } 1/2 \\ 1 & \text{prob } 1/2 \end{cases}$ X_i's are iid.

Then # of 1's seen is $W = \sum_{i=1}^{1000,000} X_i$. As X_i is Bernoulli(1/2), we know $E[X_i] = \frac{1}{2}$, $Var(X_i) = E[X_i^2] - (E[X_i])^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

Then $E[W] = 1{,}000{,}000 \underbrace{E[X_i]}_{1/2} = 500{,}000$, $Var(W) = 250{,}000 \Rightarrow \sigma_W = 500$

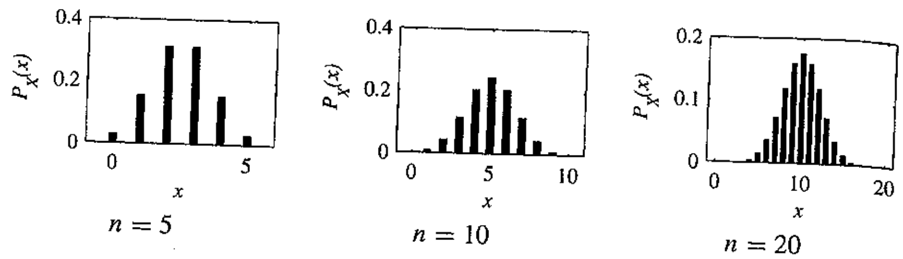

Figure 6.1 The PMF of the X, the number of heads in n coin flips for $n = 5, 10, 20$. As n increases, the PMF more closely resembles a bell-shaped curve.

We say the sum Z_n is standardized since for all n

$$E[Z_n] = 0, \qquad \text{Var}[Z_n] = 1. \tag{6.73}$$

Theorem 6.14 ***Central Limit Theorem***

Given $X_1, X_2, \ldots$, a sequence of iid random variables with expected value μ_X and variance σ_X^2, the CDF of $Z_n = (\sum_{i=1}^{n} X_i - n\mu_X)/\sqrt{n\sigma_X^2}$ has the property

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z).$$

The proof of this theorem is beyond the scope of this text. In addition to Theorem 6.14, there are other central limit theorems, each with its own statement of the sums W_n. One remarkable aspect of Theorem 6.14 and its relatives is the fact that there are no restrictions on the nature of the random variables X_i in the sum. They can be continuous, discrete, or mixed. In all cases the CDF of their sum more and more resembles a Gaussian CDF as the number of terms in the sum increases. Some versions of the central limit theorem apply to sums of sequences X_i that are not even iid.

To use the central limit theorem, we observe that we can express the iid sum $W_n = X_1 + \cdots + X_n$ as

$$W_n = \sqrt{n\sigma_X^2}\, Z_n + n\mu_X. \tag{6.74}$$

The CDF of W_n can be expressed in terms of the CDF of Z_n as

$$F_{W_n}(w) = P\left[\sqrt{n\sigma_X^2}\, Z_n + n\mu_X \le w\right] = F_{Z_n}\left(\frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right). \tag{6.75}$$

For large n, the central limit theorem says that $F_{Z_n}(z) \approx \Phi(z)$. This approximation is the basis for practical applications of the central limit theorem.

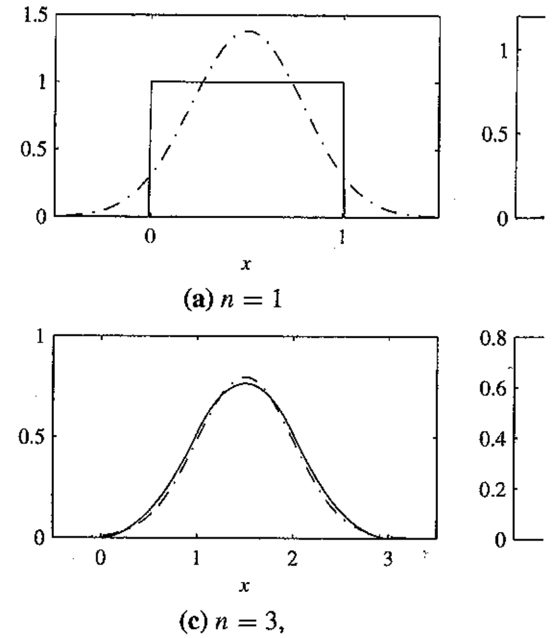

Figure 6.2 The PDF of W_n, the sum of n uniform (0, 1) central limit theorem approximation for $n = 1, 2, 3, 4$. Th while the — · — line denotes the Gaussian approximation.

Definition 6.2 ***Central Limit Theorem Approximation***

Let $W_n = X_1 + \cdots + X_n$ be the sum of n iid random $\text{Var}[X] = \sigma_X^2$. *The central limit theorem approximati*

$$F_{W_n}(w) \approx \Phi\left(\frac{w - n}{\sqrt{n\sigma}}\right.$$

We often call Definition 6.2 a Gaussian approximati

Example 6.11 To gain some intuition into the central limit thec tinuous random variables X_i, where each rand

$$W_n = X_1 + \cdots$$

Recall that $E[X] = 0.5$ and $\text{Var}[X] = 1/12$. Ther $n/2$ and variance $n/12$. The central limit theo approach a Gaussian CDF with the same exp since W_n is a continuous random variable, we would converge to a Gaussian PDF. In Figure 6 PDF of a Gaussian random variable with the san W_1 is a uniform random variable with the rect This figure also shows the PDF of W_1, a Gau

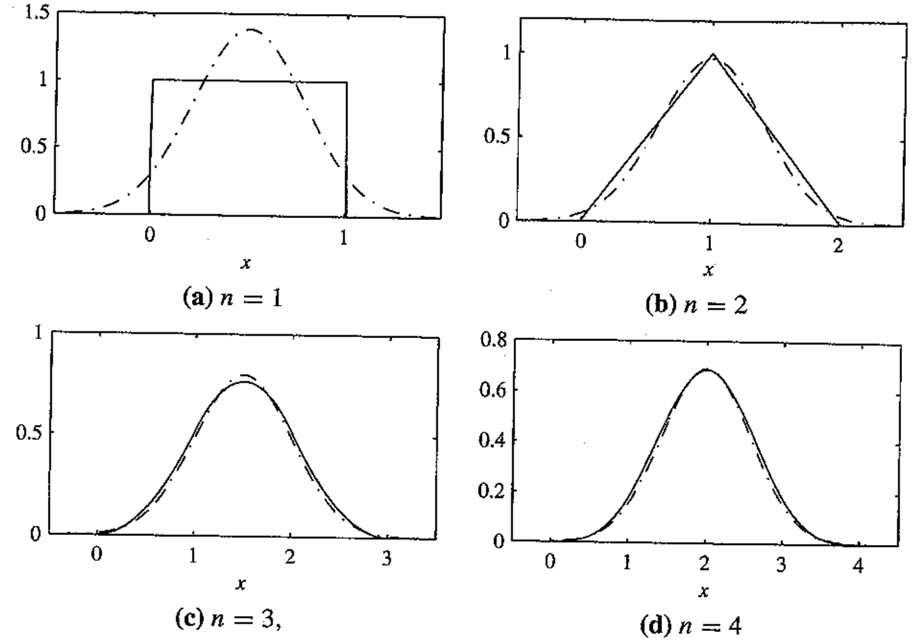

Figure 6.2 The PDF of W_n, the sum of n uniform (0, 1) random variables, and the corresponding central limit theorem approximation for $n = 1, 2, 3, 4$. The solid — line denotes the PDF $f_{W_n}(w)$, while the $- \cdot -$ line denotes the Gaussian approximation.

Definition 6.2 ***Central Limit Theorem Approximation***

Let $W_n = X_1 + \cdots + X_n$ be the sum of n iid random variables, each with $E[X] = \mu_X$ and $\mathrm{Var}[X] = \sigma_X^2$. The central limit theorem approximation to the CDF of W_n is

$$F_{W_n}(w) \approx \Phi\left(\frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right).$$

We often call Definition 6.2 a Gaussian approximation for W_n.

Example 6.11 To gain some intuition into the central limit theorem, consider a sequence of iid continuous random variables X_i, where each random variable is uniform (0,1). Let

$$W_n = X_1 + \cdots + X_n. \tag{6.76}$$

Recall that $E[X] = 0.5$ and $\mathrm{Var}[X] = 1/12$. Therefore, W_n has expected value $E[W_n] = n/2$ and variance $n/12$. The central limit theorem says that the CDF of W_n should approach a Gaussian CDF with the same expected value and variance. Moreover, since W_n is a continuous random variable, we would also expect that the PDF of W_n would converge to a Gaussian PDF. In Figure 6.2, we compare the PDF of W_n to the PDF of a Gaussian random variable with the same expected value and variance. First, W_1 is a uniform random variable with the rectangular PDF shown in Figure 6.2(a). This figure also shows the PDF of W_1, a Gaussian random variable with expected

By the central limit theorem approximation ~~theorem~~ (def. 6.2)

$$P[W \geq 502{,}000] = 1 - P[W \leq 502{,}000]$$

approx. equal

$$\approx 1 - \Phi\left(\frac{502{,}000 - \overset{E[W]}{500{,}000}}{\underset{\text{std. } W = \sqrt{Var(W)}}{500}}\right)$$

$$= Q(4) = 3.17 \times 10^{-5}$$

Moment Generating Function (MGF)

Recall from signals + systems

convolution in time $\Leftrightarrow$ multiplication in frequency.

Here, we saw that $Y = X_1 + X_2 \Rightarrow f_Y(y) = f_{X_1}(y) * f_{X_2}(y)$.

Is there an easier way than convolution to get the pdf of Y?

YES! MGF

Definition: For a RV X, the moment generating function (MGF) of X is:

$$\phi_X(s) = E_X\left[e^{sX}\right]$$

$$= \int_{-\infty}^{+\infty} e^{sx} f_X(x)\, dx. \quad \text{for } X \text{ continuous}$$

$$= \sum_{x_i \in S_X} e^{sx_i} P_X(x_i) \quad \text{for } X \text{ discrete.}$$

Random Variable	PMF or PDF	MGF $\phi_X(s)$
Bernoulli (p)	$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$	$1-p+pe^s$
Binomial (n,p)	$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$	$(1-p+pe^s)^n$
Geometric (p)	$P_X(x) = \begin{cases} p(1-p)^{x-1} & x=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$\frac{pe^s}{1-(1-p)e^s}$
Pascal (k,p)	$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$	$(\frac{pe^s}{1-(1-p)e^s})^k$
Poisson (α)	$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x=0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$e^{\alpha(e^s-1)}$
Disc. Uniform (k,l)	$P_X(x) = \begin{cases} \frac{1}{l-k+1} & x=k,k+1,\ldots,l \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{sk}-e^{s(l+1)}}{1-e^s}$
Constant (a)	$f_X(x) = \delta(x-a)$	e^{sa}
Uniform (a,b)	$f_X(x) = \begin{cases} \frac{1}{b-a} & a<x<b \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{bs}-e^{as}}{s(b-a)}$
Exponential (λ)	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{\lambda}{\lambda-s}$
Erlang (n,λ)	$f_X(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$(\frac{\lambda}{\lambda-s})^n$
Gaussian (μ,σ)	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$	$e^{s\mu+s^2\sigma^2/2}$

Table 6.1 Moment generating function for families of random variables.

includes all $s \leq 0$. Because the MGF and PMF or PDF form a transform pair, the MGF is also a complete probability model of a random variable. Given the MGF, it is possible to compute the PDF or PMF. The definition of the MGF implies that $\phi_X(0) = E[e^0] = 1$. Moreover, the derivatives of $\phi_X(s)$ evaluated at $s = 0$ are the moments of X.