Lecture 9

- **Read:** Chapter 6.6-6.8, 7.1-7.4.

Central Limit Theorem and Sample Mean

- Central Limit Theorem
- Applications of the Central Limit Theorem
- Sample Mean: Expected Value and Variance
- Sample Mean: Useful Inequalities
- Sample Mean: Sample Mean of Large Numbers
- Sample Mean: Law of Large Numbers

Sums of Random Variables

- $W_1 = X_1 + X_n$
- $f_{X_1,\ldots,X_n}(X_1, \ldots, X_n)$
- Special techniques
  - For $E[W]$ and $Var[W]$
  - When $X_1, \ldots X_n$ are iid
  - Limit theorems for large values of n

Central Limit Theorem (I)

- States that the CDF of a sum of random variables converges to a Gaussian CDF.
- Allows us to use the properties of Gaussian random variables to obtain accurate estimates of probabilities associated with sums of random variables.
  - In many cases, exact calculation of these probabilities is extremely difficult.

Central Limit Theorem (II)

- Review: $X_1, X_2, \dots$ iid Gaussian RVs
- $W_n = X_1 + \dots + X_n$ is Gaussian with

$$E[W_n] = n\mu_X$$
$$Var[W_n] = n\sigma_X^2$$

- What if $X_1, X_2, \dots$ are not Gaussian?

Sum of Bernoulli RVs

- 50 flips of a fair coin: $X_i = 1$ is H on flip i.
- W_n is binomial

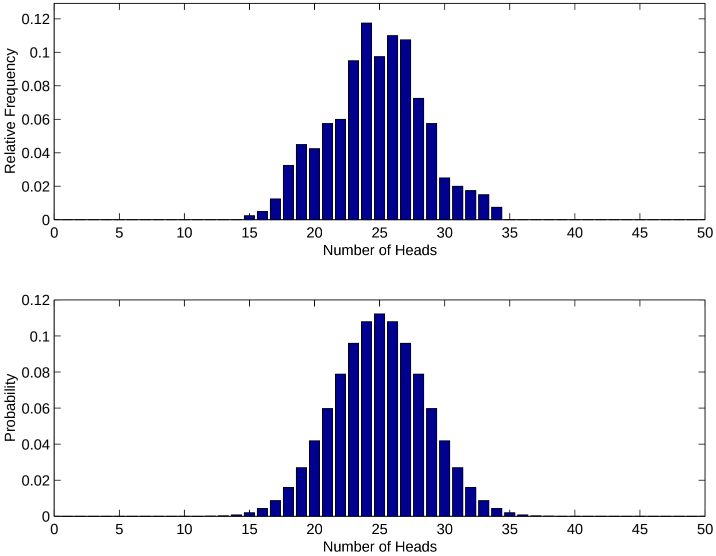
$$p_{W_n}(w) = \begin{cases} \binom{50}{w}(1/2)^{50} & \text{, w} = 0, 1, \ldots, 50 \\ 0 & \text{, otherwise} \end{cases}$$

- What does this look like?

Sum of Bernoulli RVs (cont.)

Number of heads in 50 flips of a fair coin:

400 experimental repetitions vs. the binomial PMF



Central Limit Theorem

- Suppose X_i are iid RVs and let $W_n = X_1 + X_2 + ... + X_n$.
- Define Z_n

$$Z_n = \frac{W_n - E[W_n]}{\sqrt{Var[W_n]}} = \frac{W_n - n\mu_X}{\sqrt{n} \cdot \sigma_X}$$

- As $n \to \infty$, $Z_n \sim N(0, 1)$ (Alternatively, $P[Z_n \leq z] = \Phi(z)$).

Central Limit Theorem: Proof (I)

- We want to show that the MGF of Z_n approaches the MGF of a Gaussian RV.
- Consider $\phi_{Z_n}(s) = E\left[exp\left(s\frac{W_n - n\mu_X}{\sqrt{n}\cdot\sigma_X}\right)\right]$.

$$
\begin{aligned}
\frac{W_n - n\mu_X}{\sqrt{n}\cdot\sigma_X} &= \frac{\sum X_i - n\mu_X}{\sqrt{n}\cdot\sigma_X} \\
&= \frac{\sum_{i=1}^{n}(X_i - \mu_X)}{\sqrt{n}\cdot\sigma_X} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{(X_i - \mu_X)}{\sigma_X} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \quad \text{, where,} Y_i = \frac{X_i - \mu_X}{\sigma_X}
\end{aligned}
$$

Central Limit Theorem: Proof (II)

- MGF of Z_n is $\phi_{Z_n}(s) = E\left[exp\left(s\frac{W_n - n\mu_X}{\sqrt{n}\cdot\sigma_X}\right)\right]$.
- Replacing $\frac{W_n - n\mu_X}{\sqrt{n}\cdot\sigma_X}$ with $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$, the MGF becomes

$$
\begin{aligned}
\phi_{Z_n}(s) &= E\left[exp\left(\frac{s}{\sqrt{n}}\sum_{i=1}^{n} Y_i\right)\right] \\
&= \left[\phi_Y\left(\frac{s}{\sqrt{n}}\right)\right]^n
\end{aligned}
$$

Central Limit Theorem: Proof (III)

- The Taylor series of a real function $f(x)$ that is infinitely differentiable in a neighborhood of a real number a is the power series

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots$$
$$= \sum_{0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

- Writing the Taylor series expansion for $\phi_Y\left(\frac{s}{\sqrt{n}}\right)$, as $n \to \infty$,

$$\phi_Y\left(\frac{s}{\sqrt{n}}\right) = 1 + \left.\frac{\partial \phi_Y\left(\frac{s}{\sqrt{n}}\right)}{\partial s}\right|_{s=0} \cdot \frac{s}{\sqrt{n}} + \left.\frac{\partial^2 \phi_Y\left(\frac{s}{\sqrt{n}}\right)}{\partial^2 s}\right|_{s=0} \cdot \frac{1}{2!}\left(\frac{s}{\sqrt{n}}\right)^2$$
$$= 1 + E[Y] \cdot \frac{s}{\sqrt{n}} + E[Y^2] \cdot \frac{s^2}{2n} \quad \text{where, } Y = \frac{X-\mu}{\sigma}$$

- Since $E[Y] = 0$ and $E[Y^2] = Var[Y] = 1$,

$$\phi_{Z_n}(s) \approx \left(1 + \frac{s^2}{2n}\right)^n$$

Central Limit Theorem: Proof (IV)

- **Fact: (Limit definition of the exponential function)**

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$$

- So,

$$\lim_{n\to\infty}\phi_{Z_n}(s) = e^{s^2/2}$$

- **Note:** $e^{s^2/2}$ is the MGF of a Gaussian RV with mean 0 and variance 1.
- So, the theorem has been proven.

Central Limit Theorem: Setup

- X_i iid.
- $W_n = X_1 + X_2 + ... + X_n$
- Let $Z_n = \frac{W_n - E[W_n]}{\sqrt{Var[W_n]}}$.
- Then, as $n \to \infty$, $Z_n \sim N(0, 1)$ or alternatively, $P[Z_n \leq z] = \Phi(z)$ (Alternatively, $W_n \sim N(n\mu_X, n\sigma_X^2)$.)

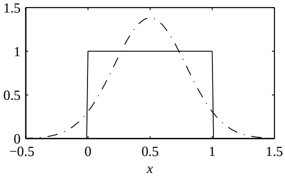
Applying the Central Limit Theorem (CLT)

- $W_n = X_1 + X_2 + ... + X_n$, X_i are iid.
- Find $P[W_n \leq w] = P\left[\frac{W_n - n\mu_X}{\sqrt{n\sigma_X^2}} \leq \frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right]$.
- **Note:** $E[W_n] = n\mu_X$ and $Var[W_n] = n\sigma_X^2$

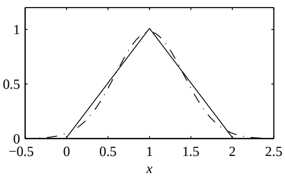
$$P[W_n \leq w] \overset{CLT}{\approx} P\left[Z \leq \frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right] \text{, where } Z \sim N(0,1)$$

$$= \Phi\left(\frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right)$$

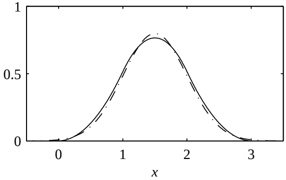
Central Limit Theorem for Uniform RVs



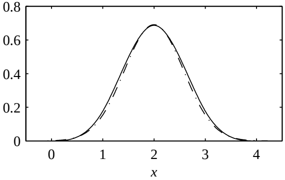
(a) $n = 1$



(b) $n = 2$

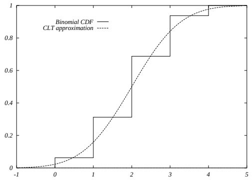


(c) $n = 3$,

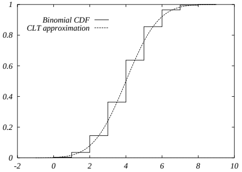


(d) $n = 4$

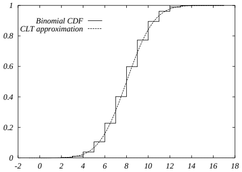
Central Limit Theorem for Binomial RVs



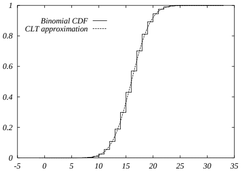
$n = 4,\ p = 1/2$



$n = 8,\ p = 1/2$



$n = 16,\ p = 1/2$



$n = 32,\ p = 1/2$

Central Limit Theorem: Example

- Suppose orders at a restaurant are iid with mean $\mu_X = 8$ TL and standard deviation $\sigma_X = 2$ TL.
- Estimate the probability that the first 100 customers spend a total exceeding 840 TL.

$$\text{Consider } W_{100} = \sum_{i=1}^{100} X_i \ , \left[\frac{W_n - n\mu_X}{\sqrt{n} \cdot \sigma_X} = \frac{W_{100} - 100 \times 8}{\sqrt{100} \cdot 2} \right]$$

- Our goal is to compute:

$$\begin{aligned}
P[W_{100} \geq 840] &= P\left[\frac{W_{100} - 800}{10 \times 2} \geq \frac{840 - 800}{10 \times 2}\right] \\
&= P\left[\frac{W_{100} - 800}{10 \times 2} \geq 2\right] \\
&= P\left[Z \geq 2\right] \quad \text{, where } Z \sim N(0,1) \\
&= 1 - \Phi(2) \\
P[Z \geq 2] &= 2.28 \times 10^{-2} \\
P[W_{100} \geq 840] &\approx 2.28 \times 10^{-2}
\end{aligned}$$

Sample Mean

- To define the sample mean, consider repeated independent trials of an experiment.
- Each trial results in one observation of a random variable, X. After n trials, we have sample values of the n RVs, $X_1, ..., X_n$ all with the same PDF as X.
- The sample mean is the numerical average of the observations.

Sample Mean

- $X_1, ..., X_n$ are iid, each with PDF $f_X(x)$
- The sample mean of X is the RV:

$$M_n(X) = \frac{X_1 + X_2 + ... + X_n}{n}$$

- Remember $M_n(X)$ is an RV!
- $M_n(X)$ is not the expected value $E[X]$.
- As n increases without bound, $M_n(X)$ predictably approaches $E[X]$.

Mean and Variance of $M_n(X)$

- **Theorem:** The sample mean $M_n(X)$ has expectation and variance:

$$E[M_n(X)] = E[X]$$
$$Var[M_n(X)] = \frac{Var[X]}{n}$$

- $\lim_{n\to\infty} Var[M_n(X)] = 0$ suggests $M_n(X) \to E[X]$.
- How does a sequence of RVs approach a constant?
  - Markov inequality
  - Chebyshev inequality
  - Chernoff inequality

Useful Inequalities

- Often, the performance of a system is determined by the probability of an undesirable event.
- The primary performance for a digital communication system is the probability of a bit error.
- For a fire alarm, the probability of a false alarm may be ignored when there is an actual fire.
- When an exact calculation is too difficult, an upper bound offers a way to guarantee that the probability of the undesirable event will not be too high.

Markov Inequality

- **Theorem: (Markov Inequality)** For nonnegative RV X and $c > 0$,

$$P[X \geq c] \leq \frac{E[X]}{c}$$

- **Proof:** Since X is a nonnegative RV, $f_X(x) = 0$ for $x < 0$.

$$\begin{aligned}
E[X] &= \int_0^{+\infty} x f_X(x) dx \\
&= \underbrace{\int_0^{c} x f_X(x) dx}_{0 \leq} + \int_c^{\infty} x f_X(x) dx \\
&\geq \int_c^{\infty} x f_X(x) dx \geq c \int_c^{\infty} f_X(x) dx = cP[X \geq c]
\end{aligned}$$

Markov Inequality: Example

- $X =$ height (in meters) of a random adult
- $E[X] = 1.60$ m
- We want to know the probability of finding an adult with height over 3.20 m.
- Markov inequality says:

$$P[X \geq 3.20] \leq \frac{1.60}{3.20} = \frac{1}{2} \longrightarrow \text{a very crude bound!}$$

- Statement is true, but is so weak it sounds wrong!

Chebyshev Inequality (I)

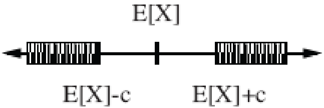
- Let $X = (Y - \mu_Y)^2$ and apply the Markov inequality:

$$P[(Y - \mu_Y)^2 \geq c^2] \leq \frac{E[(Y - \mu_Y)^2]}{c^2}$$

- **Reminder:** Markov $X \geq 0$, $c \geq 0$, $P[X \geq c^2] \leq \frac{E[X]}{c^2}$
- Now suppose we are given Y.
- Let $X = (Y - \mu_Y)^2$, note $X \geq 0$.

$$P[(Y - \mu_Y)^2 \geq c^2] \leq \frac{Var[Y]}{c^2}$$
$$P[|Y - \mu_Y| \geq c] \leq \frac{Var[Y]}{c^2}$$

Chebyshev Inequality (II)



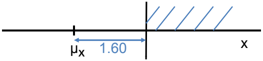
- **Theorem: (Chebyshev Inequality)** For any RV X and $c > 0$,

$$P[|X - \mu_X| \geq c] \leq \frac{Var[X]}{c^2}$$

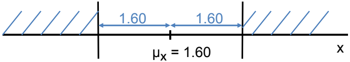
Chebyshev Inequality: Example

- For height X, $E[X] = 1.60$ m and $\sigma_X = 0.30$ m

$$P[X \geq 3.20] = P[X - \mu_X \geq 3.20 - \mu_X] = P[X - \mu_X \geq 1.60]$$



$$P[|X - \mu_X| \geq 1.60]$$



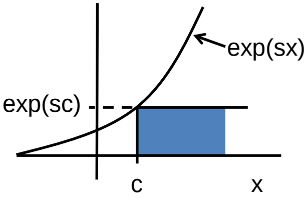
- Chebyshev: $$P[X \geq 3.20] = P[|X - \mu_X| \geq 1.60]$$
$$\leq \frac{Var[X]}{(1.60)^2}$$
$$\leq \frac{\sigma_X^2}{(1.60)^2}$$
$$\leq \frac{(0.30)^2}{(1.60)^2} \approx 0.035$$

Chernoff Bound

- **Theorem: (Chernoff Bound)** For an arbitrary RV X and a constant c,

$$P[X \geq c] \leq \min_{s \geq 0} e^{-sc} \phi_X(s)$$

Chernoff Bound: Proof 1



- **Proof:** Let $Y = exp(sX)$ for $s \geq 0$. Then, since $Y \geq 0$, by Markov's inequality

$$P[Y \geq e^{sc}] \leq \frac{E[Y]}{e^{sc}}$$

- So,

$$\begin{aligned} P[X \geq c] &\leq e^{-sc} E[exp(sX)] \\ &\leq e^{-sc} \phi_X(s) \end{aligned}$$

- Now optimize to get the **tightest** upper bound.

Chernoff Bound: Proof 2

- **Proof:** In terms of the unit step function, $u(x)$, we observe that

$$P[X \geq c] = \int_{c}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} u(x-c) f_X(x)dx$$

- For all $s \geq 0$, $u(x-c) \leq e^{s(x-c)}$. This implies

$$\begin{aligned} P[X \geq c] &= \int_{-\infty}^{\infty} e^{s(x-c)} f_X(x)dx \\ &= e^{-sc} \int_{-\infty}^{\infty} e^{sx} f_X(x)dx = e^{-sc} \phi_X(s) \end{aligned}$$

- This claim is true for any $s \geq 0$.
- Hence, the upper bound must hold when we choose s to minimize $e^{-sc}\phi_X(s)$.

Chernoff Bound

- By referring to the MGF of an RV, the Chernoff bound generally offers a better bound than the Chebyshev inequality.
- The Chernoff bound can be applied to any random variable.
- However, for small values of c, $e^{-sc}\phi_X(s)$ will be minimized by a negative value of s.
- In this case, the minimizing nonnegative $s = 0$ and the Chernoff bound gives the trivial answer: $P[X \geq c] \leq 1$.

Chernoff Bound: Example

- If the height of a randomly chosen adult is a Gaussian RV with expected value $E[X] = 1.60$ meters and standard deviation $\sigma_X = 0.30$ meters, use the Chernoff bound to find an upper bound on $P[X \geq 3.20]$.

. . . .

- Since X is $N(1.60, 0.30)$, we find in the table for MGFs that the MGF of X is
$$\phi_X(s) = e^{(3.20s+0.3^2s^2)/2}$$
- Thus, the Chernoff bound is
$$P[X \geq 3.20] \leq \min_{s\geq 0} e^{-3.20s} e^{(3.20s+0.09s^2)/2} = \min_{s\geq 0} e^{(0.09s^2-3.20s)/2}$$
- To find the minimizing s, it is sufficient to choose s to minimize $h(s) = 0.09s^2 - 3.20s$.
- Setting the derivative $dh(s)/ds = 0.18s - 3.20 = 0$ yields $s = 160/9 \approx 17.77$.
- Applying $s = 160/9$ to the bound yields
$$P[X \geq 3.20] \leq e^{(0.09s^2-3.20s)/2}\Big|_{s=160/9} = 6.65 \times 10^{-7}$$

Chebyshev Inequality Applied to $M_n(X)$ (I)

- Recall Chebyshev bound:

$$P[|X - \mu_X| \geq c] \leq \frac{Var[X]}{c^2}$$

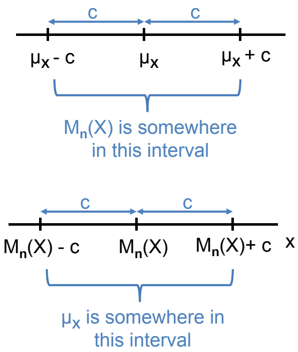
$$\begin{aligned} \text{Consider } M_n(X) &\rightarrow E[M_n(X)] = E[X] \\ &\rightarrow Var[M_n(X)] = \frac{Var[X]}{n} \end{aligned}$$

Chebyshev Inequality Applied to $M_n(X)$ (II)

- For any $c > 0$,

$$P[|M_n(X) - \mu_X| \geq c] \leq \frac{Var[X]}{nc^2} = \alpha$$

$$\begin{aligned} P[|M_n(X) - \mu_X| < c] &= 1 - P[|M_n(X) - \mu_X| \geq c] \\ &\geq 1 - \frac{Var[X]}{nc^2} = 1 - \alpha \end{aligned}$$



Chebyshev Inequality Applied to $M_n(X)$ (III)

$$P[|M_n(X) - \mu_X| < c] \geq 1 - \frac{Var[X]}{nc^2} = 1 - \alpha$$

- The probability of the sample mean being more than $\pm c$ away from the expected value is less than $\frac{Var[X]}{nc^2}$.

$$\begin{aligned} P[|M_n(X) - \mu_X| < c] &\geq 1 - \alpha \\ P[\mu_X \in [M_n(X) - c, M_n(X) + c]] &\geq 1 - \alpha \end{aligned}$$

- $c =$ size of confidence interval
- $\alpha = \frac{Var[X]}{nc^2} =$ confidence coefficient
- small $\alpha =$ high confidence

Example: Voter Survey (I)

- "Out of 1103 voters, the percentage of those that support Jones is 58% $\pm$ 3%."
- In this case, the data provides an estimate $M_n(X) = 0.58$.
- What is the confidence coefficient α of this statement?

Example: Voter Survey (II)

- **Experiment:** Observe whether a random voter support Jones
- $X = 1$ if the voter supports Jones, and $X = 0$ otherwise.
- X is a Bernoulli RV: $E[X] = p$, $Var[X] = p(1-p)$.
- Problem statement gives $\pm$ 3%. So, for $c = 0.03$

$$P[|M_n(X) - \mu_X| < 0.03] \geq 1 - \frac{p(1-p)}{n(0.03)^2} = 1 - \alpha$$

$$\text{confidence is } \alpha = \frac{p(1-p)}{n(0.03)^2}$$

- Note: $Var[X]$ is a function of p here.
- We can find an upper bound for $Var[X]$ by taking the derivative with respect to p and setting it equal to 0.

$$Var[X] = f(p) = p(1-p) = p - p^2$$

$$f'(p) = 1 - 2p = 0$$

$$p = 1/2$$

- So, p$=$ 1/2 maximizes $Var[X]$ and this maximum value is $Var[X] = 1/2 \cdot (1 - 1/2) = 1/4$.
- For all p, $Var[X] = p(1-p) \leq 1/4$.

Example: Voter Survey (III)

$$\text{confidence is } \alpha = \frac{p(1-p)}{n(0.03)^2}$$
$$\text{For all } p,\ Var[X] = p(1-p) \leq 0.25,$$
$$\alpha \leq \frac{0.25}{n(0.03)^2} = \frac{277.778}{n}$$

- So, for $n = 1103$ samples, $\alpha \leq 0.25$.
- This means that $1 - \alpha \geq 0.75$, or we are 75% confident.

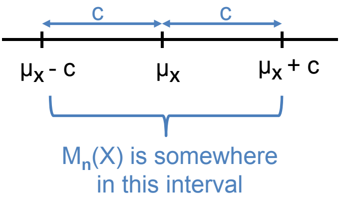
Summary

- $M_n(X) = \frac{X_1+X_2+...+X_n}{n}$: "sample mean"
  - $E[M_n(X)] = \mu_X$
  - $Var[M_n(X)] = \frac{Var[X]}{n}$
- Chebyshev Bound

$$P[|Y - \mu_Y| \geq c] \leq \frac{Var[Y]}{c^2}$$

Key Result

- $P[|M_n(X) - \mu_X| \geq c] \leq \frac{Var[X]}{nc^2} = \alpha$
- $P[|M_n(X) - \mu_X| \leq c] \geq 1 - \alpha$
- The probability that $M_n(X)$ lies within $\pm c$ of its own mean exceeds $1 - \alpha =$ our confidence level



Averaged Measurements (I)

- X_i is the ith independent measurement (in cm) of a board, the exact length of which is b cm: $X_i = b + Z_i$
- Z_i is random, $E[Z_i] = 0$, $\sigma_Z = 1$
- Use $M_n(X)$ to get accurate estimate
- By taking the average of a large number of measurements, we hope to get the estimated length $M_n(X)$ close to b.
- How many measurements should be made to guarantee that with a probability of $1 - \alpha = 0.99$ or higher, the estimate is within 0.1 cm of the exact length of the board? That is, what should n be?

$$
\begin{aligned}
M_n(X) &= \frac{X_1 + X_2 + ... + X_n}{n} \\
E[M_n(X)] &= b \\
Var[M_n(X)] &= \frac{1}{n}
\end{aligned}
$$

Averaged Measurements (II)

- $E[X_i] = b, \; Var[X_i] = Var[Z] = 1$

$$P[|M_n(X) - b| < 0.1] \geq 1 - \frac{1}{n(0.1)^2} = 1 - \frac{100}{n}$$

$$P[|M_n(X) - b| < 0.1] \geq 0.99 \quad \text{if } \frac{100}{n} \leq 0.01$$

- We need $n \geq 10,000$ measurements.

Averaged Measurements (III)

- Find n if Z_i are iid Gaussian.

$$M_n(X) = b + \frac{1}{n}(Z_1 + ... + Z_n) = b + M_n(Z) \text{ is Gaussian with}$$

$$E[M_n(Z)] = 0$$

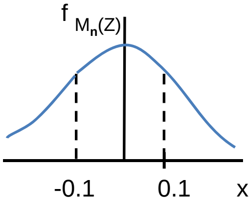
$$Var[M_n(Z)] = \frac{Var[Z]}{n} = \frac{1}{n}$$

$$\implies M_n(Z) \sim N\left(0, \frac{1}{n}\right)$$

Averaged Measurements (IV)

- Thus,

$$\begin{aligned} P[|M_n(X) - b| < 0.1] &= P[|b + M_n(Z) - b| < 0.1] \\ &= P[|M_n(Z)| < 0.1] \end{aligned}$$



Averaged Measurements (V)

- Since $M_n(Z) \sim N\left(0, \frac{1}{n}\right)$,

$$
\begin{aligned}
M_n(Z) &= \frac{1}{\sqrt{n}} Y \quad \text{, where } Y \sim N(0,1) \\
P[|M_n(Z)| < 0.1] &= P\left[\left|\frac{1}{\sqrt{n}} Y\right| < 0.1\right] \\
&= P[|Y| \le 0.1\sqrt{n}] \\
&= \Phi(0.1\sqrt{n}) - (1 - \Phi(0.1\sqrt{n})) \\
&= 2\Phi(0.1\sqrt{n}) - 1 \\
P[|M_n(X) - b| < 0.1] &\ge 0.99
\end{aligned}
$$

Averaged Measurements (VI)

- We can compute n such that:

$$
\begin{aligned}
2\Phi(0.1\sqrt{n}) - 1 &\geq 0.99 \\
\Phi(0.1\sqrt{n}) &\geq \frac{1.99}{2} \\
\Phi(0.1\sqrt{n}) &\geq 0.995 \\
0.1\sqrt{n} &\geq 2.58 \\
n &\geq 666
\end{aligned}
$$

- We find that $n \geq 666$, a number much smaller than the 10,000 we had found previously.

Weak Law of Large Numbers

- Suppose $X_1, X_2, ..., X_n$ are iid, then

$$\frac{X_1 + X_2 + ... + X_n}{n} \longrightarrow \mu_X \text{ “in probability”}$$

- That is,

$$P\left[\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu_X\right| \geq c\right] \longrightarrow 0 \text{ as } n \to \infty \text{ for any } c > 0$$

- So, the sample mean converges in probability to the true mean.
- With high probability, the sample mean for a large enough fixed value of n is close to the true mean.
- **Proof:** Use the Chebyshev bound:

$$P[|M_n(X) - \mu_X| \geq c] \leq \frac{Var[X]}{nc^2}$$

Central Limit Theorem

- Suppose $X_1, X_2, ..., X_n$ are iid, then

$$\frac{X_1 + X_2 + ... + X_n - n\mu_X}{\sqrt{n}\sigma_X} \xrightarrow{n \to \infty} N(0,1) \text{ "in distribution"}$$

- That is,

$$P\left[\frac{X_1 + X_2 + ... + X_n - n\mu_X}{\sqrt{n}\sigma_X} \leq z\right] = \Phi(z)$$

Central Limit Theorem

Example: Estimating confidence that one is close to

- $P[|S_n - \mu|] \geq \frac{\sigma\epsilon}{\sqrt{n}} \approx P[|Z \geq \epsilon|]$
- used to determine when to stop a simulation, i.e., how big should n be?

Central Limit Theorem: Applications

- Central Limit Theorem explains the common appearance of the "bell curve" in density estimates applied to real world data
- In cases like electronic noise, examination grades, etc., we can often regard a single measured value as the weighted average of a large number of small effects
- Signal processing: smoothing signals
- Traffic engineering: how many vehicles safe in tunnel and for how long?
- Statistical mechanics
- Network engineering
- Bacteria in food samples
- Many other applications...

Approximations and Closeness

Approximations: consider system models as time evolves, or some scaling parameter becomes small/large, e.g., noise small, population large.
Applications: estimation, design of circuits and system subject to noise, stochastic approximations/models for systems
Which notion of closeness is most useful? Depends!

- for closeness of most realizations then convergence in probability is appropriate, i.e., we are willing to accept a small probability of failure
- in some cases, we only care that the distributions be the same, then convergence in distribution is appropriate, e.g., convergence to the steady state.

What Limit Do I Use When?

Depends on application...

- LLN - study of stability of systems
- CLT - heavy traffic limits (study of performance when systems are overloaded)