MTH 361

Lectures 17-19

Yevgeniy Kovchegov
Oregon State University

Topics:

- Independent random variables.
- Sums of random variables.
- Law of large numbers (LLN).
- Central limit theorem (CLT).
- Examples.

Independent random variables.

Definition. Random variables X and Y are said to be **independent** if the events

$$\{\ a \le X \le b\ \} \quad \text{and} \quad \{\ c \le Y \le d\ \}$$

are independent, for any a, b, c, and d. Namely,

$$P(\ a \le X \le b \quad \text{and} \quad c \le Y \le d\ ) = P(a \le X \le b) \cdot P(c \le Y \le d)$$

Properties:

- If X and Y are both discrete random variables. They are independent if and only if

$$P(X = a \quad \text{and} \quad Y = b) \ = \ P(X = a) \cdot P(Y = b)$$

for all a and b in the corresponding sample spaces.

- If X and Y are independent then

$$E[f(X) \cdot g(Y)] = E[f(X)] \cdot E[g(Y)]$$

for any pair of functions, f and g.

Sums of independent random variables.

• If X and Y are independent discrete random variables, with probability mass functions $p_{\mathrm{x}}(x) = P(X = x)$ and $p_{\mathrm{y}}(y) = P(Y = y)$, then their sum, $Z = X + Y$ is also a discrete random variable with probability mass function

$$p_{\mathrm{z}}(a) = \sum_{x,y:\; x+y=a} p_{\mathrm{x}}(x) \cdot p_{\mathrm{y}}(y)$$

which can be rewritten as a **convolution sum**:

$$p_{\mathrm{z}}(a) = \sum_{x} p_{\mathrm{x}}(x) \cdot p_{\mathrm{y}}(a-x)$$

• If X and Y are independent continuous random variables, with density functions f_{x} and f_{y}, then $Z = X + Y$ is also a continuous random variable with its density f_Z given as a **convolution integral**,

$$f_{\mathrm{z}}(a) = \int_{-\infty}^{\infty} f_{\mathrm{x}}(x) \cdot f_{\mathrm{y}}(a-x)\; dx$$

Sums of independent random variables.

$$p_{x+y}(a) \;=\; \sum_x p_x(x)\cdot p_y(a-x)$$

• **Example.** Let X be binomial with parameters (n_1, p) and Y be binomial with parameters (n_2, p). Then their probability mass functions are

$$p_x(k) = \binom{n_1}{k}\cdot p^k \cdot (1-p)^{n_1-k} \qquad \text{for } k = 0, 1, \ldots, n_1$$

and

$$p_y(k) = \binom{n_2}{k}\cdot p^k \cdot (1-p)^{n_2-k} \qquad \text{for } k = 0, 1, \ldots, n_2$$

If X and Y are independent, then their sum will have the following distribution: for $j = 0, 1, \ldots, n_1 + n_2$,

$$p_{x+y}(j) \;= \sum_k p_x(k)p_y(j-k) = \sum_{\substack{0\le k\le n_1\\ j-n_2\le k\le j}} \binom{n_1}{k}\cdot p^k\cdot(1-p)^{n_1-k}\binom{n_2}{j-k}\cdot p^{j-k}\cdot(1-p)^{n_2-j+k}$$

$$= \; p^j\cdot(1-p)^{n_1+n_2-j}\cdot \sum_{\substack{0\le k\le n_1\\ j-n_2\le k\le j}} \binom{n_1}{k}\binom{n_2}{j-k}$$

Sums of independent random variables.

$$p_{x+y}(a) = \sum_x p_x(x) \cdot p_y(a-x)$$

- **Example (continued).** Let X be binomial with parameters (n_1, p) and Y be binomial with parameters (n_2, p). If X and Y are independent, then their sum will have the following distribution: for $j = 0, 1, \ldots, n_1 + n_2$,

$$p_{x+y}(j) = \sum_k p_x(k)\, p_y(j-k) = p^j\,(1-p)^{n_1+n_2-j} \sum_{\substack{0 \le k \le n_1 \\ j-n_2 \le k \le j}} \binom{n_1}{k}\binom{n_2}{j-k},$$

where $\sum_{\substack{0 \le k \le n_1 \\ j-n_2 \le k \le j}} \binom{n_1}{k}\binom{n_2}{j-k} = \binom{n_1+n_2}{j}$ since both sides represent the number of $n_1 + n_2$ long strings one can make with j A's and $n_1 + n_2 - j$ B's.

Hence,

$$p_{x+y}(j) = \binom{n_1+n_2}{j} \cdot p^j \cdot (1-p)^{n_1+n_2-j} \qquad \text{for } j = 0, 1, \ldots, n_1 + n_2$$

Thus $X + Y$ is binomial with parameters $(n_1 + n_2, p)$.

Sums of independent random variables.

$$p_{x+y}(a) \;=\; \sum_x p_x(x) \cdot p_y(a-x)$$

• **Example.** Let X be Poisson with parameter λ_1 and Y be Poisson with parameter λ_2. Then their probability mass functions are

$$p_x(k) = e^{-\lambda_1} \cdot \frac{\lambda_1^k}{k!} \qquad \text{for } k = 0, 1, \ldots$$

and

$$p_y(k) = e^{-\lambda_2} \cdot \frac{\lambda_2^k}{k!} \qquad \text{for } k = 0, 1, \ldots$$

If X and Y are independent, then their sum will have the following distribution: for $n = 0, 1, \ldots,$

$$p_{x+y}(n) \;=\; \sum_{k=0}^{\infty} p_x(k) \cdot p_y(n-k) \;=\; \sum_{k=0}^{n} e^{-\lambda_1} \cdot \frac{\lambda_1^k}{k!}\; e^{-\lambda_2} \cdot \frac{\lambda_2^{n-k}}{(n-k)!}$$

$$=\; e^{-(\lambda_1+\lambda_2)} \cdot \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \cdot \lambda_1^k \cdot \lambda_2^{n-k}$$

Sums of independent random variables.

$$p_{x+y}(a) \;=\; \sum_x p_x(x) \cdot p_y(a-x)$$

• **Example (continued).** Let X be Poisson with parameter λ_1 and Y be Poisson with parameter λ_2. If X and Y are independent, then their sum will have the following distribution: for $n = 0, 1, \ldots,$

$$p_{x+y}(n) \;=\; e^{-(\lambda_1+\lambda_2)} \cdot \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \cdot \lambda_1^k \cdot \lambda_2^{n-k} \;=\; \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \cdot \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \cdot \lambda_1^k \cdot \lambda_2^{n-k}$$

$$=\; \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \cdot \sum_{k=0}^{n} \binom{n}{k} \cdot \lambda_1^k \cdot \lambda_2^{n-k} \;=\; e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^n}{n!}$$

by the Binomial Theorem.

Hence, $X+Y$ is Poisson with parameter $\lambda_1+\lambda_2$.

Sums of independent random variables.

$$f_{x+y}(a) = \int_{-\infty}^{\infty} f_x(x) \cdot f_y(a-x)\ dx$$

• **Example.** Let X and Y each be uniform over $[0,1]$. Then each will be distributed according to the following probability density function:

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

If X and Y are independent, then their sum will have the following continuous distribution:

$$f_{x+y}(a) = \int_{-\infty}^{\infty} f(x) \cdot f(a-x)\ dx = \int_{0}^{1} f(a-x)dx$$

Observe that $\int_{0}^{1} f(a-x)dx = 0$ whenever $a < 0$ or $a > 2$.

There are two more cases: $0 \leq a \leq 1$ and $1 \leq a \leq 2$.

Sums of independent random variables.

- **Example (continued).** $f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$

If $0 \le a \le 1$,

$$f_{x+y}(a) = \int_0^1 f(a-x)dx = \int_0^a dx = a$$

Now, if $1 \le a \le 2$,

$$f_{x+y}(a) = \int_0^1 f(a-x)dx = \int_{a-1}^1 dx = 2 - a$$

Therefore,

$$f_{x+y}(a) = \begin{cases} a & \text{if } 0 \le a \le 1 \\ 2-a & \text{if } 1 \le a \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Properties of convolutions.

Convolution of functions f and g is defined

$$f \circ g(a) = \int_{-\infty}^{\infty} f(x) \cdot g(a - x)\ dx$$

- Convolution is commutative: $f \circ g(a) = g \circ f(a)$

- If you are familiar with Fourier transforms: $\widehat{f \circ g}(\xi) = \widehat{g}(\xi)\widehat{f}(\xi)$

Sums of random variables.

• **Theorem.** Expectation of a sum of random variables is equal to the sum of expectations.

$$E[X+Y] = E[X] + E[Y]$$

Notice: we don't require X and Y to be independent here. The above is true even if they are dependent random variables.

• **Theorem.** If X and Y are independent random variables then,

$$Var(X+Y) = Var(X) + Var(Y)$$

Here we need independence.

• **Example.** Let X be Poisson with parameter λ_1 and Y be Poisson with parameter λ_2. For the case when X and Y are independent, we proved that $X+Y$ is Poisson with parameter $\lambda_1+\lambda_2$. Then

$$E[X] + E[Y] = \lambda_1 + \lambda_2 = E[X+Y]$$

$$Var(X) + Var(Y) = \lambda_1 + \lambda_2 = Var(X+Y)$$

Sums of random variables.

• **Example.** Let X and Y each be uniform over $[0,1]$. For the case when X and Y are independent, we proved that $X+Y$ is distributed according to

$$f_{\mathrm{x+y}}(a) = \begin{cases} a & \text{if } 0 \leq a \leq 1 \\ 2-a & \text{if } 1 \leq a \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Check that

$$E[X+Y] = E[X] + E[Y]$$

and

$$Var(X) + Var(Y) = Var(X+Y)$$

Sums of random variables.

• **Example (continued).** Let X and Y each be uniform over $[0,1]$. For the case when X and Y are independent, we proved that $X+Y$ is distributed according to

$$f_{x+y}(a) = \begin{cases} a & \text{if } 0 \le a \le 1 \\ 2-a & \text{if } 1 \le a \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Here $E[X] = E[Y] = \int_0^1 x dx = \frac{1}{2}$ while

$$E[X+Y] = \int_{-\infty}^{\infty} x f_{x+y}(x) dx = \int_0^1 x^2 dx + \int_1^2 x(2-x) dx$$

$$= \frac{1}{3} + \left[x^2 - \frac{x^3}{3}\right]_1^2 = \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3} = 1 = E[X] + E[Y]$$

Sums of random variables.

• **Example (continued).** Let X and Y each be uniform over $[0,1]$. For the case when X and Y are independent, we proved that $X+Y$ is distributed according to

$$f_{x+y}(a) = \begin{cases} a & \text{if } 0 \le a \le 1 \\ 2-a & \text{if } 1 \le a \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Here $Var(X) = E[X^2] - \Big(E[X]\Big)^2 = \int_0^1 x^2 dx - \frac{1}{4} = \frac{1}{12} = Var(Y)$

while $E[X+Y] = 1$, and therefore

$$Var(X+Y) = E[(X+Y)^2] - 1 = \int_{-\infty}^{\infty} x^2 f_{x+y}(x) dx - 1 = \int_0^1 x^3 dx + \int_1^2 x^2(2-x) dx - 1$$

$$= \frac{1}{4} + \left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_1^2 - 1 = \frac{1}{4} + \frac{16}{3} - 4 - \frac{2}{3} + \frac{1}{4} - 1 = \frac{1}{6} = Var(X) + Var(Y)$$

Sums of independent random variables.

Consider n **independent identically distributed** (i.i.d.) random variables

$$X_1,\ X_2, \ldots,\ X_n$$

There

$$E[X_1] = E[X_2] = \ldots = E[X_n] = \mu$$

and

$$Var(X_1) = Var(X_2) = \ldots = Var(X_n) = \sigma^2$$

Suppose μ and σ are finite. Then

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] = n\mu$$

and

$$Var(X_1 + X_2 + \ldots + X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n) = n\sigma^2$$

Thus

$$X_1 + X_2 + \ldots + X_n = n\mu \pm \sqrt{n}\sigma$$

and

$$\frac{X_1 + X_2 + \ldots + X_n}{n} = \mu \pm \frac{\sigma}{\sqrt{n}}$$

Sums of independent random variables.

Consider n **independent identically distributed** (i.i.d.) random variables

$$X_1,\ X_2, \ldots,\ X_n$$

There

$E[X_1 + X_2 + \ldots + X_n] = n\mu$ and $Var(X_1 + X_2 + \ldots + X_n) = n\sigma^2$

Thus

$$X_1 + X_2 + \ldots + X_n = n\mu \pm \sqrt{n}\sigma$$

and

$$\frac{X_1 + X_2 + \ldots + X_n}{n} = \mu \pm \frac{\sigma}{\sqrt{n}}$$

- **Law of Large Numbers.** Given $\epsilon > 0$, then

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) = 0$$

In other words, for n large enough, $\frac{X_1+X_2+\ldots+X_n}{n} \approx \mu$.

Law of Large Numbers.

Consider n **independent identically distributed** (i.i.d.) random variables $X_1, X_2, \ldots, X_n$. There $E[X_1 + X_2 + \ldots + X_n] = n\mu$ and $Var(X_1 + X_2 + \ldots + X_n) = n\sigma^2$

- **Law of Large Numbers.** Given $\epsilon > 0$, then

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) = 0$$

In other words, for n large enough, $\frac{X_1+X_2+\ldots+X_n}{n} \approx \mu$.

Proof: Recall the Chebyshev inequality: for any $\kappa > 0$,

$$P\Big(\Big|X - E[X]\Big| \geq \kappa\Big) \;\leq\; \frac{Var(X)}{\kappa^2}$$

Applying it to $X_1 + X_2 + \ldots + X_n$, we obtain

$$P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \;=\; P\Big(|X_1+X_2+\ldots+X_n-n\mu| > n\epsilon\Big)$$

$$\leq \frac{Var(X_1 + X_2 + \ldots + X_n)}{n^2\epsilon^2} \;=\; \frac{n\sigma^2}{n^2\epsilon^2} \;=\; \frac{\sigma^2}{n\epsilon^2} \longrightarrow 0 \quad \text{as } n \to \infty$$

Central Limit Theorem.

Consider n **independent identically distributed** (i.i.d.) random variables $X_1,\ X_2, \ldots,\ X_n$, each with mean μ and variance σ^2. There

$E[X_1 + X_2 + \ldots + X_n] = n\mu$ and $Var(X_1 + X_2 + \ldots + X_n) = n\sigma^2$

and therefore $X_1 + X_2 + \ldots + X_n = n\mu \pm \sqrt{n}\sigma$.

Hence, $\frac{X_1+X_2+\ldots+X_n-n\mu}{\sqrt{n}\sigma} = 0 \pm 1$.

In fact, as n gets larger, $\frac{X_1+X_2+\ldots+X_n-n\mu}{\sqrt{n}\sigma}$ is distributed more and more like the standard normal random variable.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \le \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \leq \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

- **Example.** Given $0 \leq p \leq 1$. We say that X is a **Bernoulli random variable** with parameter p if

$$P(X = 1) = p \quad \text{and} \quad P(X = 0) = 1 - p$$

Think of one coin toss. There $E[X] = p$ and $Var(X) = p(1-p)$.

Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables with the same parameter p. Then

$$S_n = X_1 + X_2 + \ldots + X_n$$

is a Binomial random variable with parameters (n, p). Then

$$P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) \approx \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \leq \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Let $X_1,\ X_2, \ldots,\ X_n$ be independent Bernoulli random variables with the same parameter p. Then

$$S_n = X_1 + X_2 + \ldots + X_n$$

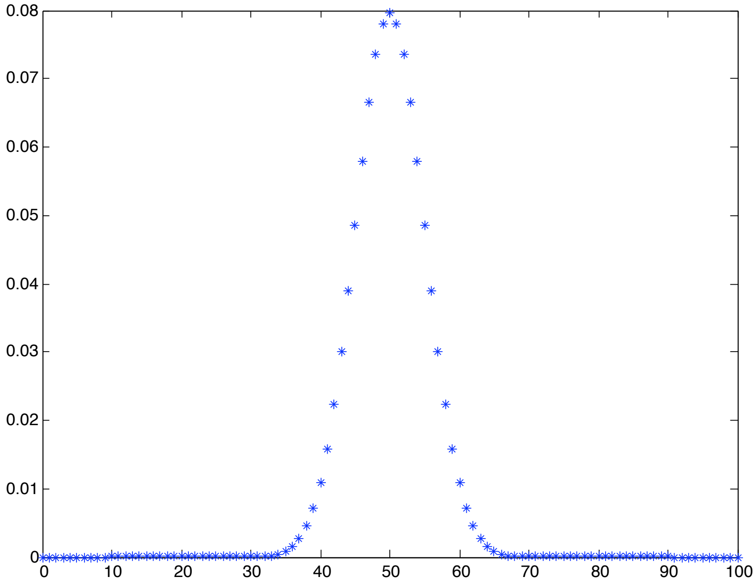
is a Binomial random variable with parameters (n, p). Then, S_n satisfies the following version of the Central Limit Theorem:

- **De Moivre - Laplace Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

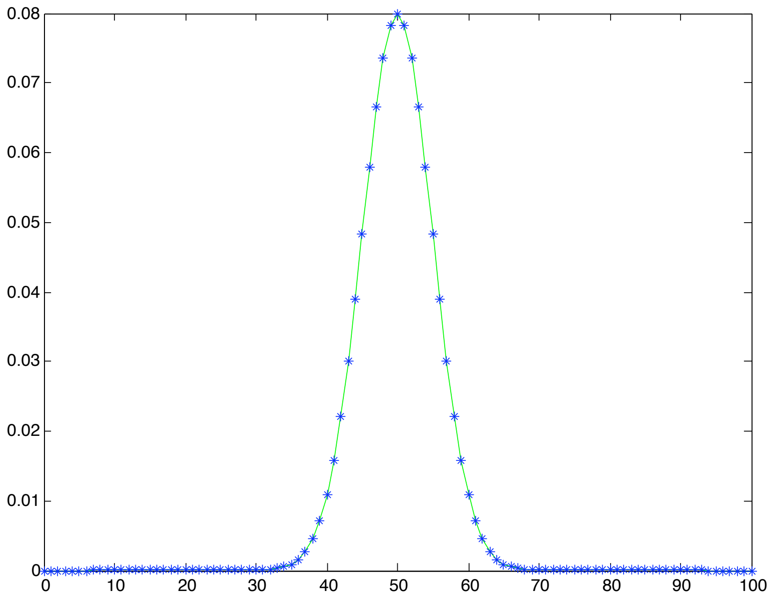
Example. Let S_n be a Binomial random variable with $n = 100$ and $p = \frac{1}{2}$. Estimate $P(S_n = 53)$ and $P(52 \leq S_n \leq 57)$.



Central Limit Theorem.

Example (continued). Observe that

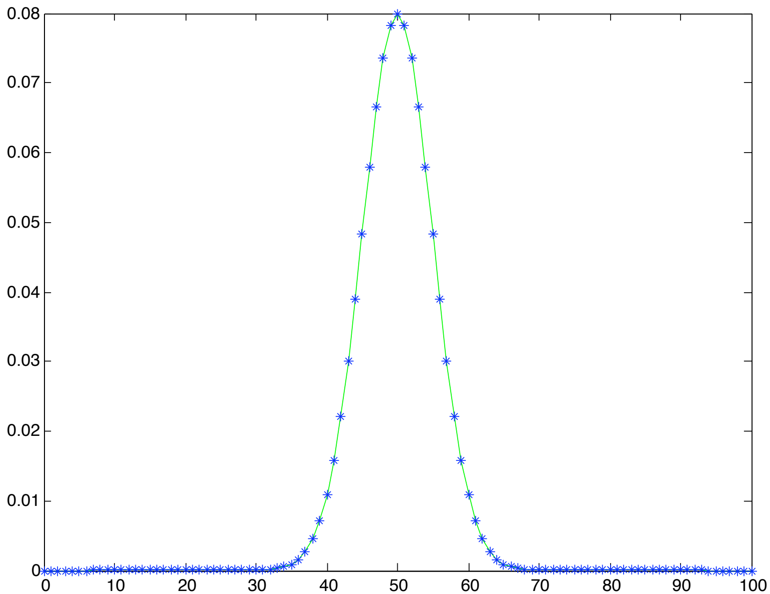
$$P(S_n = 53) = P(52.5 \leq S_n \leq 53.5)$$



Central Limit Theorem.

Example (continued). $E[S_n] = np = 50$ and

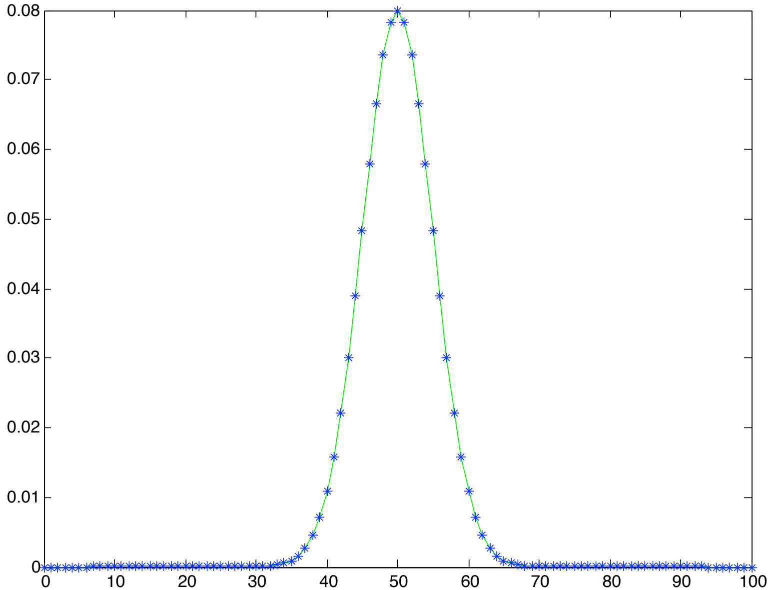
$$P(S_n = 53) = P(52.5 \leq S_n \leq 53.5) = P(2.5 \leq S_n - np \leq 3.5)$$



Central Limit Theorem.

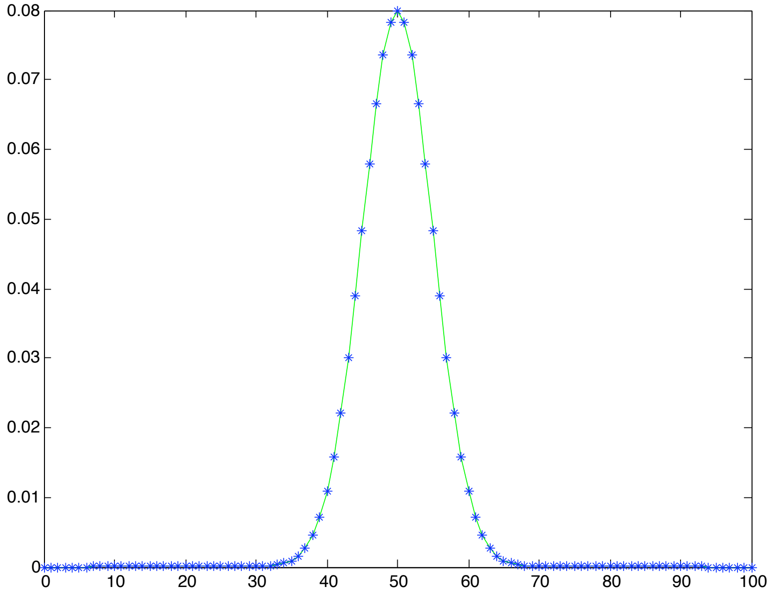
Example (continued). $Var(S_n) = np(1-p) = 25$ and

$$P(S_n = 53) = P(2.5 \leq S_n - np \leq 3.5) = P\left(\frac{2.5}{5} \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{3.5}{5}\right)$$



Central Limit Theorem.

$$P(S_n = 53) = P\left(0.5 \le \frac{S_n - np}{\sqrt{np(1-p)}} \le 0.7\right) \approx \int_{0.5}^{0.7} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$



Central Limit Theorem.

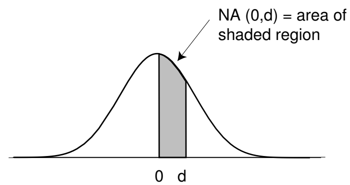
$$P(S_n = 53) = P\left(0.5 \le \frac{S_n - np}{\sqrt{np(1-p)}} \le 0.7\right) \approx \int_{0.5}^{0.7} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

$$= \int_{0}^{0.7} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx - \int_{0}^{0.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = 0.2580\ldots - 0.1915\ldots = 0.0665\ldots$$

499

Appendix A

Normal distribution table

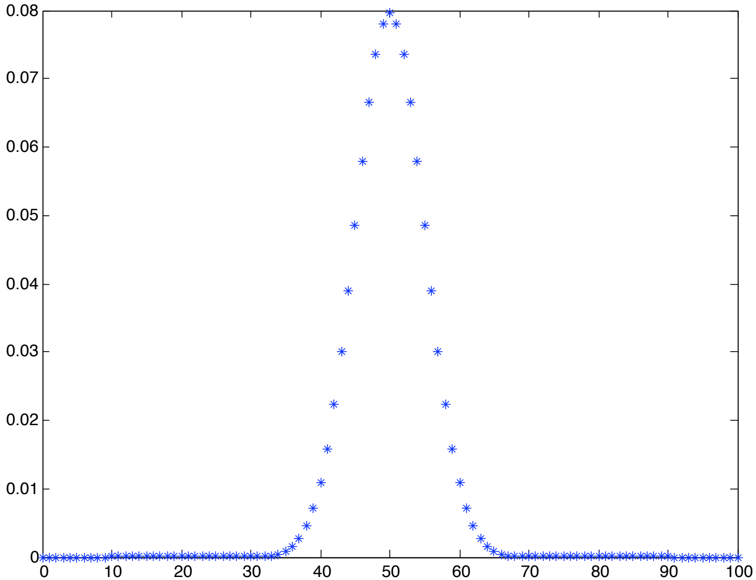


	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830

Central Limit Theorem.

Example (continued). So $P(S_n = 53) \approx$ 0.0665 . . .

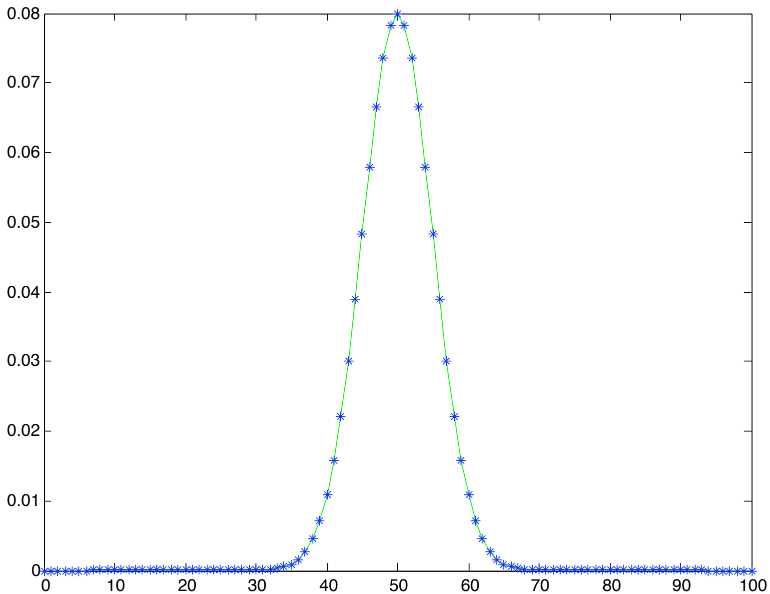
Next, estimate $P(52 \leq S_n \leq 57)$.



Central Limit Theorem.

Example (continued). Observe that

$$P(52 \leq S_n \leq 57) = P(51.5 \leq S_n \leq 57.5)$$



Central Limit Theorem.

Example (continued). Recall that $E[S_n] = np = 50$ and $Var(S_n) = np(1-p) = 25$. Thus

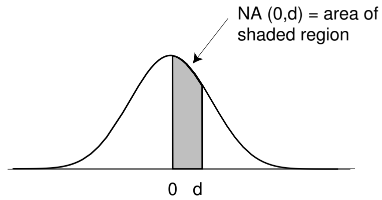
$$P(52 \le S_n \le 57) = P(51.5 \le S_n \le 57.5) = P(1.5 \le S_n - np \le 7.5)$$

$$= P\left(\frac{1.5}{5} \le \frac{S_n - np}{\sqrt{np(1-p)}} \le \frac{7.5}{5}\right) = P\left(0.3 \le \frac{S_n - np}{\sqrt{np(1-p)}} \le 1.5\right)$$

$$\approx \int_{0.3}^{1.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = \int_{0}^{1.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx - \int_{0}^{0.3} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

$$= 0.4332\ldots - 0.1179\ldots = 0.3153\ldots$$

Normal distribution table



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Central Limit Theorem.

Consider n **independent identically distributed** (i.i.d.) random variables $X_1, X_2, \ldots, X_n$, each with mean μ and variance σ^2. There

$E[X_1 + X_2 + \ldots + X_n] = n\mu$ and $Var(X_1 + X_2 + \ldots + X_n) = n\sigma^2$

and therefore $X_1 + X_2 + \ldots + X_n = n\mu \pm \sqrt{n}\sigma$.

Hence, $\frac{X_1+X_2+\ldots+X_n-n\mu}{\sqrt{n}\sigma} = 0 \pm 1$.

In fact, as n gets larger, $\frac{X_1+X_2+\ldots+X_n-n\mu}{\sqrt{n}\sigma}$ is distributed more and more like the standard normal random variable.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \leq \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \le \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

- **Example.** Given $0 \le p \le 1$. We say that X is a **Bernoulli random variable** with parameter p if

$$P(X = 1) = p \quad \text{and} \quad P(X = 0) = 1 - p$$

Think of one coin toss. There $E[X] = p$ and $Var(X) = p(1-p)$.

Let $X_1,\ X_2, \ldots,\ X_n$ be independent Bernoulli random variables with the same parameter p. Then

$$S_n = X_1 + X_2 + \ldots + X_n$$

is a Binomial random variable with parameters (n, p). Then

$$P\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) \approx \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

- **Central Limit Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \le \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables with the same parameter p. Then

$$S_n = X_1 + X_2 + \ldots + X_n$$

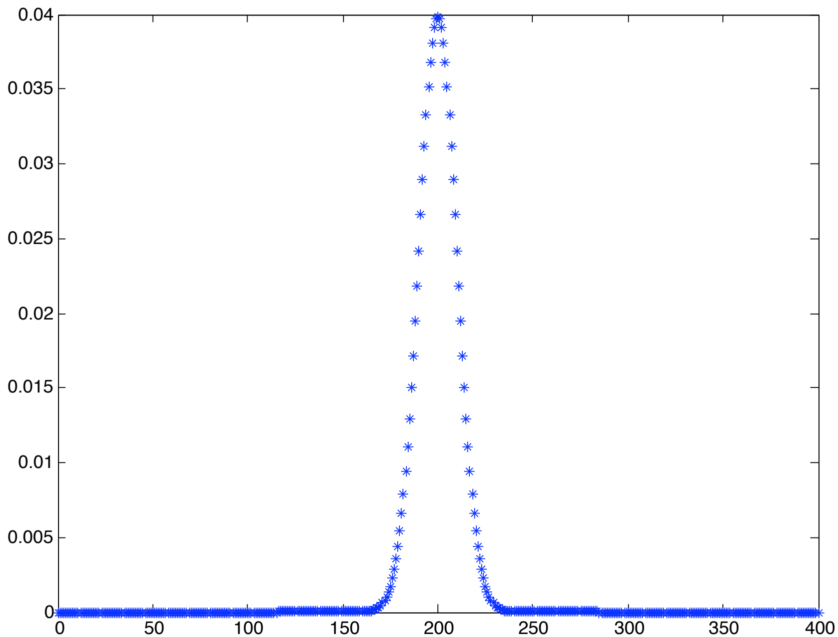
is a Binomial random variable with parameters (n, p). Then, S_n satisfies the following version of the Central Limit Theorem:

- **De Moivre - Laplace Theorem.** Given $a < b$, then

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

Central Limit Theorem.

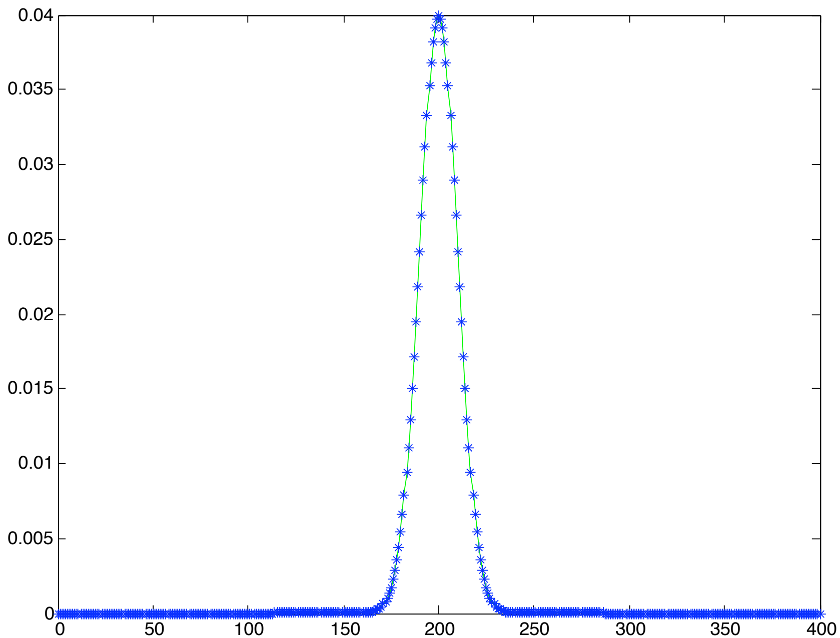
Example. Let S_n be a Binomial random variable with $n = 400$ and $p = \frac{1}{2}$. Estimate $P(S_n = 199)$ and $P(S_n \geq 192)$.



Central Limit Theorem.

Example (continued). Observe that

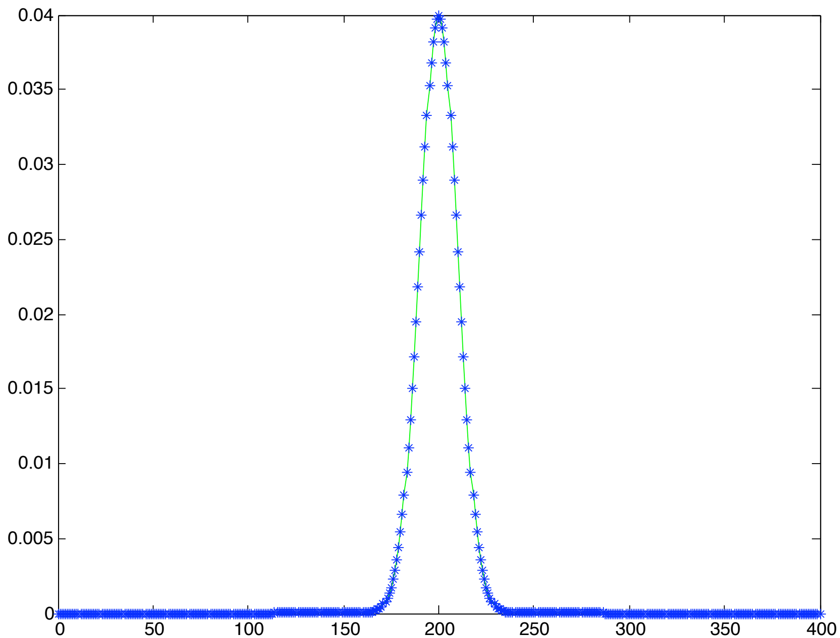
$$P(S_n = 199) = P(198.5 \leq S_n \leq 199.5)$$



Central Limit Theorem.

Example (continued). $E[S_n] = np = 200$ and

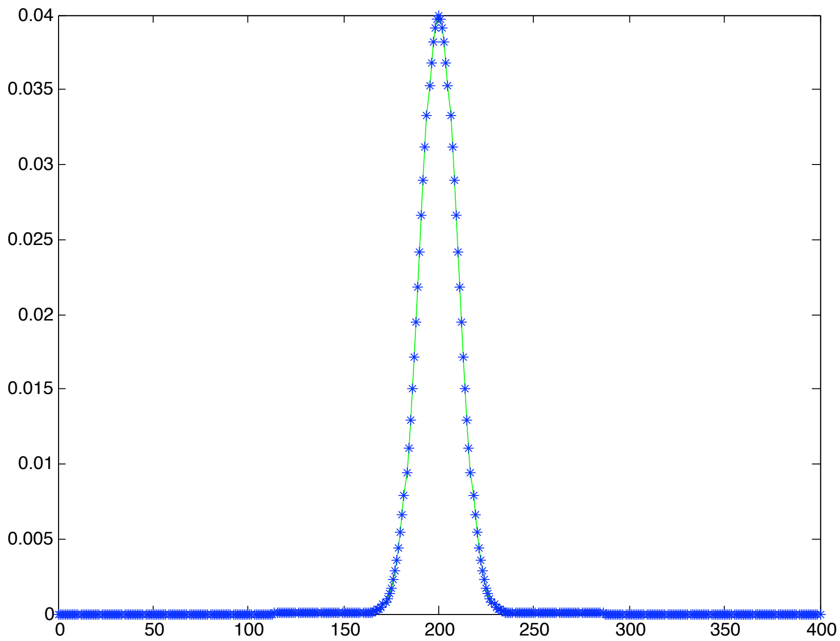
$P(S_n = 199) = P(198.5 \leq S_n \leq 199.5) = P(-1.5 \leq S_n - np \leq -0.5)$



Central Limit Theorem.

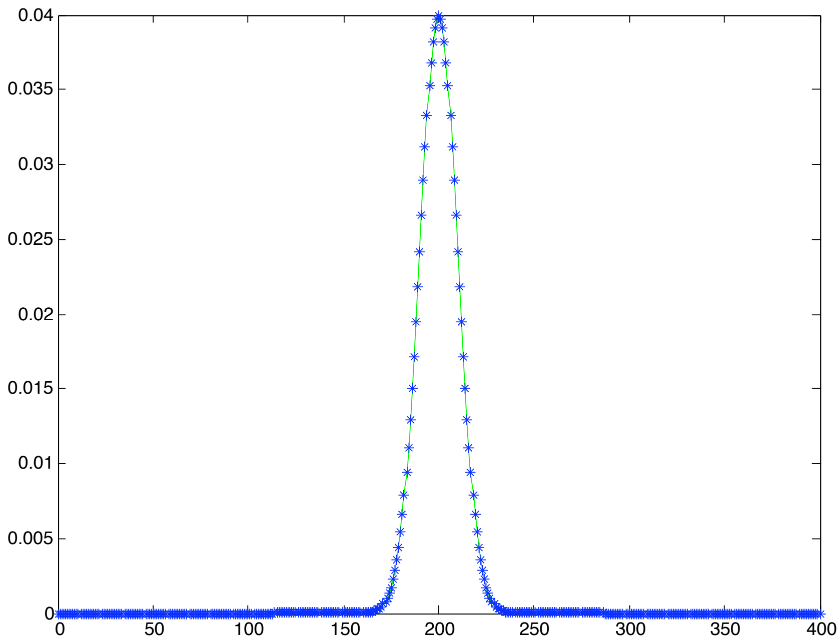
Example (continued). $Var(S_n) = np(1-p) = 100$ and

$$P(S_n = 199) = P(-1.5 \leq S_n - np \leq -0.5) = P\left(-0.15 \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq -0.05\right)$$



Central Limit Theorem.

$$P(S_n = 199) = P\left(-0.15 \le \frac{S_n - np}{\sqrt{np(1-p)}} \le -0.05\right) \approx \int_{-0.15}^{-0.05} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$



Central Limit Theorem.

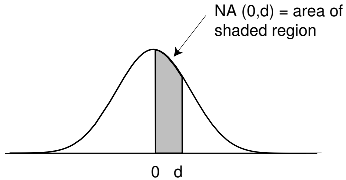
$$P(S_n = 199) = P\left(-0.15 \le \frac{S_n - np}{\sqrt{np(1-p)}} \le -0.05\right) \approx \int_{-0.15}^{-0.05} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

$$= \int_0^{0.15} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx - \int_0^{0.05} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = 0.0596\ldots - 0.0199\ldots \; = 0.0397\ldots$$

499

Appendix A

Normal distribution table

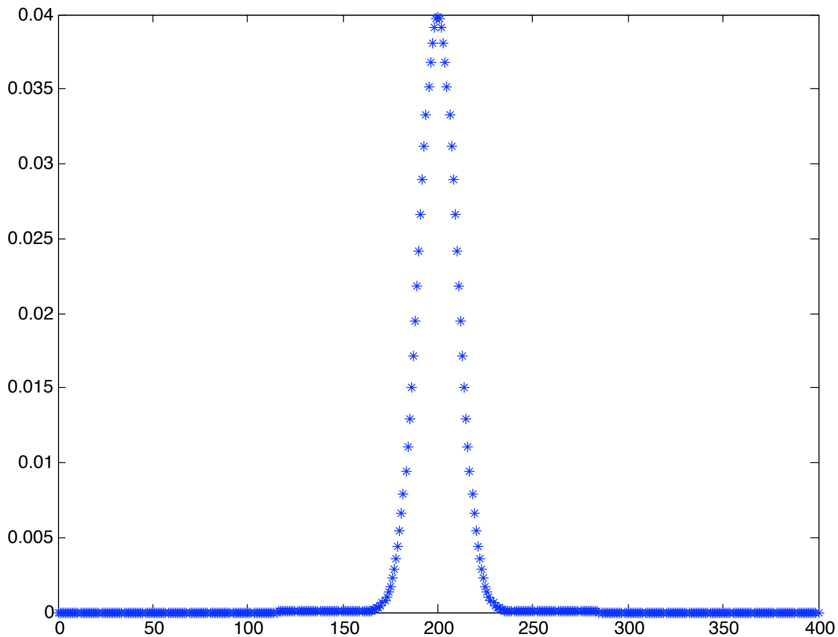


	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830

Central Limit Theorem.

Example (continued). So $P(S_n = 199) \approx 0.0397\ldots$

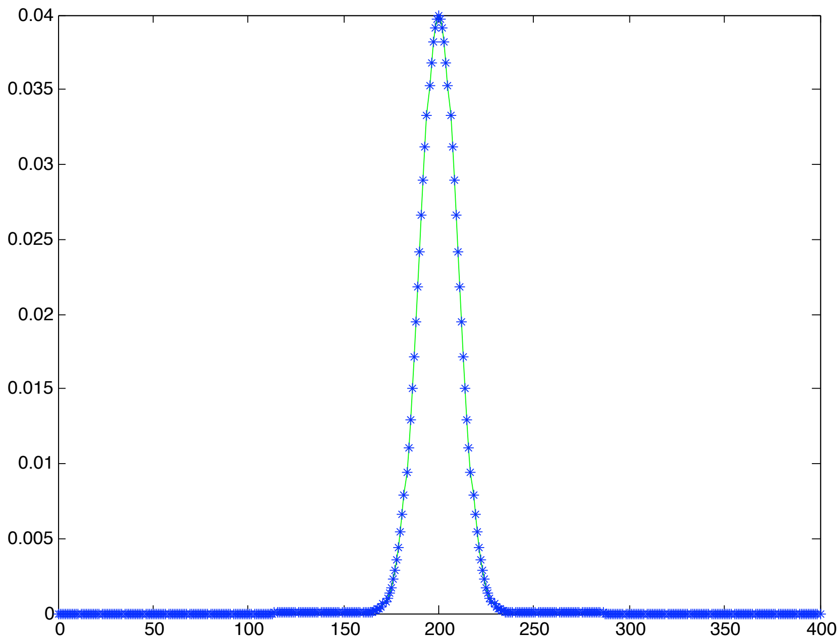
Next, estimate $P(S_n \geq 192)$.



Central Limit Theorem.

Example (continued). Observe that

$$P(S_n \geq 192) = P(191.5 \leq S_n < \infty)$$



Central Limit Theorem.

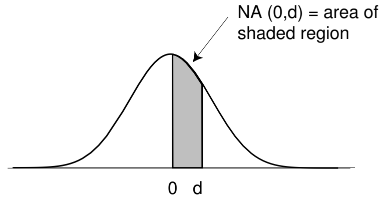
Example (continued). Recall that $E[S_n] = np = 200$ and $Var(S_n) = np(1-p) = 100$. Thus

$$P(S_n \geq 192) = P(191.5 \leq S_n < \infty) = P(-8.5 \leq S_n - np < \infty)$$

$$= P\left(-0.85 \leq \frac{S_n - np}{\sqrt{np(1-p)}} < \infty\right) \approx \int_{-0.85}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

$$= \int_{0}^{0.85} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx + \frac{1}{2} = 0.3023\ldots + 0.5 = 0.8023\ldots$$

Normal distribution table



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000