Example: Suppose that \( Z \) has standard normal distribution.
(1) Find the probability
\[
P(-0.564 \leq Z \leq 1.22).
\]
(2) Find \( z \) such that
\[
P(-0.7 \leq Z \leq z) = 0.4.
\]
Solution:

(1)
\[
P(-0.564 \leq Z \leq 1.22)
= P(Z \leq 1.22) - P(Z \leq -0.564)
= 0.8888 - 0.2877
= 0.6011
\]

(2)
\[
P(-0.7 \leq Z \leq z) = P(Z < z) - P(Z < -0.7)
= P(Z < z) - 0.242
= 0.4
\]

and
\[
P(Z < z) = 0.642
\]

Thus \( z = 0.365 \).

Many sequences of discrete distributions have normal distributions as limiting distributions. (See graphs).

Question: Consider the binomial experiment with \( n = 10,000 \) and the probability of success \( p = 0.3 \). Let \( X \) be the number of successes in the \( 10,000 \) trials. Find the approximate probability \( P(X \leq 6000) \).

Approaching Idea: We know the distribution table of binomial:
The distribution table of r.v $X$ with $n = 10,000$

<table>
<thead>
<tr>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>$\cdots$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>$C_0^n p^0 q^n$</td>
<td>$C_1^n p^1 q^{n-1}$</td>
<td>$C_2^n p^2 q^{n-2}$</td>
<td>$\cdots$</td>
<td>$C_n^n p^n q^0$</td>
</tr>
</tbody>
</table>

We have

$$P(X \leq 6000) = \sum_{i=0}^{6000} \frac{n!}{i!(n-i)!} p^i q^{n-i}$$

This is equal to the area of the corresponding histogram between 0 and 6000 (See corresponding graph). Do the histogram top broken lines converge to a curve as $n \to \infty$? If the corresponding area under the curve is easy to be calculated, this would be a wonderful idea. See the following histograms:
Unnormalized R.F. of \( B(n=20,p=0.3) \)
Unnormalized R.F. of B(n=100, p=0.3)

Change in categories (2.0)
The means keep moving as $n$ changes. This is not convenient. Therefore, instead of considering $X$ itself, we consider a normalized random variable

$$Y = \frac{X - np}{\sqrt{npq}}$$

Then, we have following histograms of $Y$. 
Normalized R.F. of $B(n=20, p=0.3)$

Change in categories (0.7)
In the case of r.v. \( Y \), its mean is always equal to zero and its variance is always equal to 1. Its histograms converge to a curve with following function

\[
f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}
\]

Then it is easy to find any area under this curve, above 0, and between any two numbers \( a < b \) from the normal distribution table (at the end of your textbook).

We have following theorem.

**Theorem:** (Central Limit Theorem) Let \( X_1, X_2, \cdots \) be an infinite sequence of independent rv’s, each with the same distribution. Suppose that the mean \( \mu \) and the variance \( \sigma^2 \) of any \( X_i \) are both finite. Then, for any numbers \( a \) and \( b \),

\[
\lim_{n \to \infty} \mathbb{P}(a \leq \sum_{i=1}^{n} X_i - n\mu \leq b) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-x^2/2} dx
\]

**Example:** Let \( X \) be a rv with binomial distribution \( b(n, p) \). Define

\[
X_i = \begin{cases} 
1 & \text{if observe a head, with prob. } p \\
0 & \text{if observe a tail, with prob. } 1 - p
\end{cases}
\]

Then, \( \mathbb{E}(X_i) = np \) and \( \text{Var}(X_i) = p(1 - p) \) for \( i = 1, \cdots, n \) and \( X = \sum_{i=1}^{n} X_i \). By Central Limit Theorem, for any two constants \( a < b \), we have

\[
\lim_{n \to \infty} \mathbb{P}(a \leq \frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1 - p)}} \leq b) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-x^2/2} dx
\]

**Applications of Normal Approximation to Binomial Dist.:**
Suppose that \( X \) is a binomial r.v. with distribution \( b(n, p) \). If the
following condition (1) is satisfied, then the following formulae can be used to find the corresponding probabilities.

Procedure:
(1) Check conditions: \( np > 5 \) and \( nq > 5 \).
(2) Find \( \mu = np \), \( p \) is the success probability and \( q = 1 - p \). \( \sigma = \sqrt{npq} \).
(3) Let

\[
Z = \frac{X - np}{\sqrt{npq}}
\]

Then \( Z \) is an approximate standard normal r.v.. For given constant integers \( a, b \), use following formulae to find approximate probabilities:

(I) \[
P(X \geq a) = P(X \geq a - 0.5) = P(X - np \geq a - 0.5 - np)
\]
\[
= P\left( \frac{X - np}{\sqrt{npq}} \geq \frac{a - 0.5 - np}{\sqrt{npq}} \right)
\]
\[
= P(Z \geq \frac{a - 0.5 - np}{\sqrt{npq}})
\]

(II) \[
P(X \leq b) = P(X \leq b + 0.5) = P(X - np \leq b + 0.5 - np)
\]
\[
= P\left( \frac{X - np}{\sqrt{npq}} \leq \frac{b + 0.5 - np}{\sqrt{npq}} \right)
\]
\[
= P(Z \leq \frac{b + 0.5 - np}{\sqrt{npq}})
\]

(III) \[
P(a \leq X \leq b) = P(a - 0.5 \leq X \leq b + 0.5)
\]
\[
= P(a - 0.5 - np \leq X - np \leq b + 0.5 - np)
\]
\[
= P\left( \frac{a - 0.5 - np}{\sqrt{npq}} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{b + 0.5 - np}{\sqrt{npq}} \right)
\]
\[
= P\left( \frac{a - 0.5 - np}{\sqrt{npq}} \leq Z \leq \frac{b + 0.5 - np}{\sqrt{npq}} \right)
\]
Example 1  Consider the binomial experiment with \( n = 10,000 \) and the probability of success \( p = 0.3 \). Let \( X \) be the number of successes in the 10,000 trials. Find the approximate probability \( \mathbb{P}(3000 \leq X \leq 3100) \).

Solution:  
(I) \( np = 10,000(0.3) = 3,000 > 5 \) and \( nq = 10,000(0.7) = 7,000 > 5 \).

(II) \( \mu = np = 3,000, \sigma = \sqrt{npq} = \sqrt{10,000(0.3)(0.7)} = 45.82 \)

(III) To find the probability \( \mathbb{P}(3000 \leq X \leq 3,100) \), according to (III), we have

\[
\mathbb{P}(a \leq X \leq b) = \mathbb{P}\left( \frac{a - 0.5 - np}{\sqrt{npq}} \leq Z \leq \frac{b + 0.5 - np}{\sqrt{npq}} \right)
\]

\[
\mathbb{P}(3000 \leq X \leq 3,100) = \mathbb{P}\left( \frac{3,000 - 0.5 - 3,000}{45.82} \leq Z \leq \frac{3,100 + 0.5 - 3,000}{45.82} \right)
\]

\[
= \mathbb{P}(-0.01 \leq Z \leq 2.19)
\]

0.4861