CHAPTER I

INTRODUCTION

Let X_j, $j = 1, 2, \ldots, n$ be independent random variables such that for some $p, q > 0$, $p + q = 1$, we have

$$P(X_j = 0) = q,$$

$$P(X_j = 1) = p$$

for all $j = 1, 2, \ldots, n$.

Let

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

It is well-known that $P(k_1 \leq S_n \leq k_2)$ can be approximated by

$$\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt$$

where

$$x_1 = \frac{k_1 - np}{\sqrt{npq}} \quad \text{and} \quad x_2 = \frac{k_2 - np}{\sqrt{npq}} .$$

In 1945, W.Feller [1] proved that

$$\left| P(k_1 \leq S_n \leq k_2) - \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt \right| < \frac{c}{\sqrt{n}}$$

where c is a constant.

In 1979, J.V. Uspensky [2] gave a better approximation by introduction a correction term in the approximation and made some adjustment to the normal probability. Limits of integration used in Uspensky's approximation are

$$T_1 = \frac{k_1 - np - \frac{1}{2}}{\sqrt{npq}} ,$$

$$T_2 = \frac{k_2 - np + \frac{1}{2}}{\sqrt{npq}} .$$

The correction term introduced is

$$\frac{q - p}{6\sqrt{2\pi npq}} \left[(1 - T_2^2) e^{-\frac{T_2^2}{2}} - (1 - T_1^2) e^{-\frac{T_1^2}{2}} \right] .$$

Uspensky's approximation can be stated as follows:

$$P(k_1 \leq S_n \leq k_2) = \frac{1}{\sqrt{2\pi}} \int_{T_1}^{T_2} e^{-\frac{t^2}{2}} dt + \frac{(q-p)}{6\sqrt{2 npq}} \left[(1 - T_2^2) e^{-\frac{T_2^2}{2}} - (1 - T_1^2) e^{-\frac{T_1^2}{2}} \right]$$

$$+ \Delta$$

where the error of approximation Δ satisfies

$$|\Delta| < \frac{c}{n}$$

for some constant c.

In this study, we generalize Upensky's result to the case where X_j, $j = 1, 2, \ldots, n$ are independent integral-valued random variables. Our main result is given in Theorem 3.11. This theorem is specialized to the case of identically distributed random variables in Chapter IV.