Probability and Random Process (SWE3026)

Central Limit Theorems

JinYeong Bak
jy.bak@skku.edu
College of Computing, SKKU

H. Pishro-Nik, "Introduction to probability, statistics, and random processes", available at https://www.probabilitycourse.com, Kappa Research LLC, 2014.

Summary of Probability Bounds

Markov's Inequality

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{EX}{a}, \quad \text{for any } a > 0.$$

Summary of Probability Bounds

Chebyshev's Inequality

For any random variable X , with $EX = \mu$ and $\mathrm{Var}(X) = \sigma^2$, we have

$$P\big(|X-\mu| \geq \epsilon\big) \leq \frac{Var(X)}{\epsilon^2}.$$

$$||$$

$$P(\underbrace{\mu-\epsilon}_{a} < X < \underbrace{\mu+\epsilon}_{b})$$

Law of Large Numbers

Definition. For i.i.d. random variables $X_1, X_2, \dots, X_n$ with $EX_i = \mu_i$ and $\mathrm{Var}(X_i) = \sigma_i^2$, the sample mean, denoted by $\overline{X}$, is defined as

$$\overline{X} = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Law of Large Numbers

The sample mean, $\overline{X}$, is also a random variable, then we have

$$
\begin{aligned}
E[\overline{X}] &= \frac{EX_1 + EX_2 + \dots + EX_n}{n} && \text{(by linearity of expectation)} \\
&= \frac{nEX}{n} && \text{(since } EX_i = EX\text{)} \\
&= EX.
\end{aligned}
$$

Law of Large Numbers

The variance of $\overline{X}$ is given by

$$\begin{aligned}
\mathrm{Var}(\overline{X}) &= \frac{\mathrm{Var}(X_1 + X_2 + \ldots + X_n)}{n^2} && (\mathrm{Var}(aX) = a^2\mathrm{Var}(X)) \\
&= \frac{\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n)}{n^2} && (X_i\text{'s are independent}) \\
&= \frac{n\mathrm{Var}(X)}{n^2} && (\mathrm{Var}(X_i) = \mathrm{Var}(X)) \\
&= \frac{\mathrm{Var}(X)}{n} = \frac{\sigma^2}{n}.
\end{aligned}$$

Law of Large Numbers

The weak law of large numbers (WLLN)

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with a finite expected value $EX_i = \mu < \infty$. Then, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\overline{X} - \mu| \geq \epsilon) = 0.$$

Law of Large Numbers

Proof:

We assume $\mathrm{Var}(X) = \sigma^2$ is finite. In this case we can use Chebyshev's inequality to write

$$P(|\overline{X} - \mu| \geq \epsilon) \leq \frac{\mathrm{Var}(\overline{X})}{\epsilon^2}$$

$$= \frac{\mathrm{Var}(X)}{n\epsilon^2},$$

which goes to zero as $n \to \infty$.

Central Limit Theorem

Note:

If $EX = \mu,\ \mathrm{Var}(X) = \sigma^2$ **and the normalized random variable is defined:**

$$Z = \frac{X - \mu}{\sigma},$$

then,

$$EZ = 0,\ \mathrm{Var}(Z) = 1.$$

Central Limit Theorem

Proof:

$$\boldsymbol{EZ = \frac{EX - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0,}$$

$$\boldsymbol{\mathrm{Var}(Z) = \frac{\mathrm{Var}(X)}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1.}$$

Central Limit Theorem

The Central Limit Theorem (CLT)

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $EX_i = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then,

$$Z_n = \frac{\overline{X} - E\overline{X}}{\sqrt{\mathrm{Var}(\overline{X})}} = \frac{\sum_{i=1}^{n} X_i/n - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma},$$

Central Limit Theorem

Converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n\to\infty} P(Z_n \leq x) = \Phi(x), \qquad \text{for all } x \in \mathbb{R},$$

Where $\Phi(x)$ is the standard normal CDF.

Note: This true regardless of the distribution of $\boldsymbol{X}$.

Central Limit Theorem

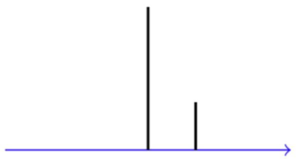
Assumptions:

- $X_1, X_2 \ldots$ are iid Bernoulli(p).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - np}{\sqrt{np(1-p)}}$.

We choose $p = \frac{1}{3}$.

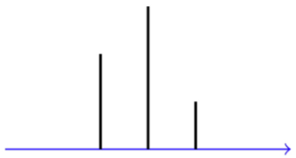
$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1



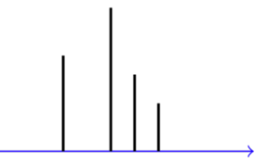
$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2



$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

PMF of Z_3



$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$

PMF of Z_{30}

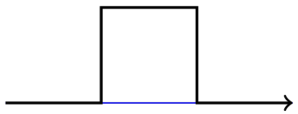
Central Limit Theorem

Assumptions:

- $X_1, X_2 \ldots$ are iid Uniform(0,1).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$.

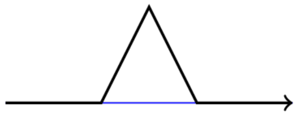
$$Z_1 = \frac{X_1 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$

PDF of Z_1



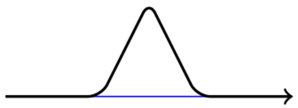
$$Z_2 = \frac{X_1 + X_2 - 1}{\sqrt{\frac{2}{12}}}$$

PDF of Z_2



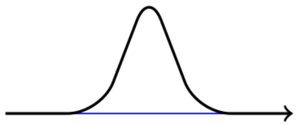
$$Z_3 = \frac{X_1 + X_2 + X_3 - \frac{3}{2}}{\sqrt{\frac{3}{12}}}$$

PDF of Z_3



$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - \frac{30}{2}}{\sqrt{\frac{30}{12}}}$$

PDF of Z_{30}



Central Limit Theorem

In practice n finite, but we still can approximate $Y = X_1 + X_2 + \cdots + X_n$ by a Normal random variable.

Central Limit Theorem

Two steps to solve problems using CLT:

$$Y = X_1, X_2, \cdots, X_n, \qquad X_i \text{ i.i.d.}$$

a) Find $EY = \mu_Y = \sum_{i=1}^{n} EX_i$, and $\mathrm{Var}(Y) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$.

b) Use $Y_n \sim N\left(\mu_{Y_n}, \mathrm{Var}(Y_n)\right)$, so we can use Φ function.

Central Limit Theorem

$$\boldsymbol{Y = X_1 + X_2 + \cdots + X_n}$$

$$\boldsymbol{EX_i = \mu,\ \mathrm{Var}(X_i) = \sigma^2}$$

$$\boldsymbol{EY = n\mu,\ \ \mathrm{Var}(Y) = n\sigma^2 \ \xrightarrow{CLT}\ Y \sim N(n\mu, n\sigma^2)}$$

Central Limit Theorem

Example.

In a digital communication system $n = 1000$ bits are transmitted over a wireless channel. Each bit will be received in error with probability $P_e = 0.1$ (error probability) independently from other bits. Let W be the number of error.

Find $P(W > 120)$