Sums of Random Variables

Dep. EE, KPU

2020년 2학기

Sums of Random Variables

- Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables
- Let S_n be their sum:

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

Sum of Random Variables

- Regardless of statistical dependence, the expected value of a sum of n random variables is equal to the sum of the expected values

$$E[X \pm Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[X \pm Y]\, f_{X,Y}(x,y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X\, f_{X,Y}(x,y)dxdy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Y\, f_{X,Y}(x,y)dxdy$$

$$= E[X] \pm E[Y]$$

generalize

$$\Rightarrow \quad E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Sum of Random Variables

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{COV}(X,Y)$$

$$Var[X+Y] = E[(X+Y-E[X]-E[Y])^2] = \mathrm{E}[\{(X-E[X]) + (Y-E[Y])\}^2]$$

$$= E[(X-E[X])^2 + (Y-E[Y])^2 + (X-E[X])(Y-E[Y]) + (Y-E[Y])(X-E[X])]$$

$$= Var[X] + Var[Y] + COV(X,Y) + COV(Y,X)$$

$$= Var[X] + Var[Y] + 2COV(X,Y)$$

Sum of Random Variables

generalize

$$\mathrm{Var}[X_1 + X_2 + \cdots + X_n\ ] = E\left\{\sum_{j=1}^{n}\left(X_j - E[X_j]\right)\sum_{k=1}^{n}(X_k - E[X_k])\right\}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{n} E[\left(X_j - E[X_j]\right)(X_k - E[X_k])]$$

$$= \sum_{k=1}^{n} Var[X_k] + \sum_{\substack{j=1 \\ j\neq k}}^{n}\sum_{k=1}^{n} COV(X_j, X_k)$$

- If $X_1, X_2, \dots, X_n$ are independent random variables, then $COV\left(X_j, X_k\right) = 0$ for $j \neq k$ and

$$\mathrm{Var}[X_1 + X_2 + \cdots + X_n\ ] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \cdots + \mathrm{Var}[X_n]$$

Example

- Find the mean and variance of sum of n independent, identically distributed (i.i.d) random variables, each with mean μ and variance σ^2

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

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

$$Var[S_n] = Var[X_1 + X_2 + \cdots + X_n] = \sum_{k=1}^{n} Var[X_k] = nVar[X_k] = n\sigma^2$$

Moment Generating Function (Recall)

- For a RV X, the moment generating function (MGF) of X is

$$\Phi_X(t) = \mathrm{E}[e^{tX}]$$

$$\Phi_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$$ Continuous

$$\Phi_X(t) = \sum_k e^{tx_k} p_X(x_k)$$ Discrete

MGF of Sum of RVs

- Let $X_1, X_2, \dots, X_n$ be independent random variables
- Their MGFs are $M_{X_1}(t), M_{X_2}(t), \dots, M_{X_n}(t)$
- If $Y = X_1 + X_2 + \cdots + X_n$, then the MGF of Y is

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t) \cdots M_{X_n}(t)$$

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t)$$

MGF of Sum of RVs

- Proof:

$$M_Y(t) = E(e^{tY}) = E\left[e^{t(X_1+X_2+\cdots+X_n)}\right]$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{t(X_1+X_2+\cdots+X_n)} f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)dx_1 dx_2 \cdots dx_n$$

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}\left(x_1\right)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

$$M_Y(t) = \int_{-\infty}^{\infty} e^{tx_1} f_{X_1}(x_1)dx_1 \int_{-\infty}^{\infty} e^{tx_2} f_{X_2}(x_2)dx_2 \cdots \int_{-\infty}^{\infty} e^{tx_n} f_{X_n}(x_n)dx_n$$

$$= M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t)$$

Example

χ^2 distribution with k DoF

$$f_X(x) = \frac{x^{(k-2)/2} e^{-x/2}}{2^{k/2}\Gamma(k/2)}, \qquad x > 0$$

$$M_X(t) = (1-2t)^{-k/2}$$

- $X_1, X_2, \dots, X_n$: independent RVs

  ~ Chi-square distribution with DoFs v_1, v_2,..., v_n

- $Y = X_1 + X_2 + \cdots + X_n$ is χ^2 distribution with DoF $v = v_1 + v_2 + \cdots + v_n$

  - Proof:

$$M_Y(t) = M_{X_1}(t)\, M_{X_2}(t) \dots M_{X_n}(t)$$

$$M_{X_i}(t) = (1-2t)^{-v_i/2}, \qquad i = 1,2,\dots,n$$

$$M_Y(t) = (1-2t)^{-v_1/2}(1-2t)^{-v_2/2} \dots (1-2t)^{-v_n/2} = (1-2t)^{-(v_1+v_2+\cdots+v_n)/2}$$

Characteristic Functions (Recall)

$$\mathrm{E}[e^{tX}] = \mathrm{E}\left[1 + tX + \frac{(tX)^2}{2!} + \cdots + \frac{(tX)^n}{n!} + \cdots\right]$$

$$= 1 + t\mathrm{E}[X] + \frac{t^2}{2!}\mathrm{E}[X^2] + \cdots + \frac{t^n}{n!}\mathrm{E}[X^n] + \cdots$$

$t = j\omega$

$$\mathrm{E}\left[e^{j\omega X}\right] = 1 + j\omega\mathrm{E}[X] + \frac{(j\omega)^2}{2!}\mathrm{E}[X^2] + \cdots$$

- The characteristic function(CF) of a RV X is defined by

Continuous $$\Phi_X(\omega) = \mathrm{E}\left[e^{j\omega X}\right] = \int_{-\infty}^{\infty} f_X(x)\, e^{j\omega x} dx$$

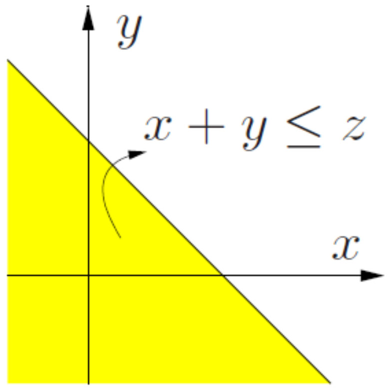
Discrete $$\Phi_X(\omega) = \mathrm{E}\left[e^{j\omega X}\right] = \sum_k e^{j\omega x_k}\, p_X(x_k)$$

- $\Phi_X(\omega)$ is obtained from the MGF when $t = j\omega$ (앞의 MGF에 $t = j\omega$ 대입한 것과 같음)
- $\Phi_X(\omega)$ is simply the Fourier transform(단, 지수승부호반대) of PDF or PMF of X
- Every pdf and its characteristic function form a unique Fourier pair:

$$\Phi_X(\omega) \Leftrightarrow f_X(x)$$

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_X(\omega)\, e^{-j\omega x} d\omega$$

PDF of Sum of Independent RVs (Recall)



- Let $Z = X + Y$

$$P(Z \leq z) = P(X + Y \leq z)$$

Integrate the joint pdf f_{XY} over the yellow region

CDF of Z: $F_Z(z) = P(Z \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_{X,Y}(x, y)\, dydx$

PDF of Z: $f_Z(z) = \frac{dF_Z(z)}{dz} = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)\, dx$

when X, Y are independent, the pdf of Z has a form of convolution:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\, dx = f_X(x) * f_Y(y)$$

PDF of Sum of Independent RVs

- Let X and Y be independent RVs and define

$$Z = X + Y$$

- The characteristic function of X is the product of characteristic functions of X and Y:

$$\Phi_Z(\omega) = \Phi_X(\omega)\Phi_Y(\omega)$$

Proof) $$\Phi_Z(\omega) = E\left[e^{j\omega Z}\right] = E\left[e^{j\omega(X+Y)}\right]$$

$$= E\left[e^{j\omega X}\right]E\left[e^{j\omega Y}\right] \quad (X \text{ and } Y \text{ are independent})$$

$$= \Phi_X(\omega)\Phi_Y(\omega)$$

$$\Phi_Z(\omega) = F\{f_Z(z)\} = F\{f_X(x) * f_Y(y)\} = \Phi_X(\omega)\Phi_Y(\omega)$$

PDF of sums of independent RVs

- Consider the sum of n independent RVs

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

- The characteristic function of S_n is

$$\begin{aligned}\Phi_S(\omega) = E\left[e^{j\omega S_n}\right] &= E\left[e^{j\omega(X_1+X_2+\cdots+X_n)}\right] \\ &= E\left[e^{j\omega X_1}\right]\cdots E\left[e^{j\omega X_n}\right] \\ &= \Phi_{X_1}(\omega)\cdots\Phi_{X_n}(\omega)\end{aligned}$$

- The pdf of S_n is found by finding the inverse Fourier of $\Phi_S(\omega)$:

$$f_S(X) = \mathcal{F}^{-1}\left[\Phi_{X_1}(\omega)\cdots\Phi_{X_n}(\omega)\right]$$

Example

- Let S_n be the sum of n independent Gaussian RVs

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

$$E[X_i] = m_i \ , \qquad Var[X_i] = \sigma_i^2 \qquad \text{for all} \quad i = 0,1,\dots,n$$

- Find the PDF of S_n

$$\Phi_{S_n}(\omega) = E\left[e^{j\omega S_n}\right] = E\left[e^{j\omega(X_1+X_2+\cdots+X_n)}\right] = \Phi_{X_1}(\omega)\cdots\Phi_{X_n}(\omega)$$

$$\Phi_{X_k}(\omega) = \int_{-\infty}^{\infty} f_{X_k}(x_k)\, e^{j\omega x_k} dx_k = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x_k-m_k)^2}{2\sigma_k^2}} e^{j\omega x_k}\, dx_k$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x_k-m_k)^2 - j\omega x_k 2\sigma_k^2}{2\sigma_k^2}} dx_k$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{x_k^2 - 2m_k x_k + m_k^2 - j\omega x_k 2\sigma_k^2}{2\sigma_k^2}} dx_k$$

Example

$$\Phi_{X_k}(\omega) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{x_k^2 - 2m_k x_k + m_k^2 - j\omega x_k 2\sigma_k^2}{2\sigma_k^2}} dx_k$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{x_k^2 - 2(m_k + j\omega\sigma_k^2)x_k + (m_k + j\omega\sigma_k^2)^2 - (m_k + j\omega\sigma_k^2)^2 + m_k^2}{2\sigma_k^2}} dx_k$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{\left(x_k - (m_k + j\omega\sigma_k^2)\right)^2 - 2j\omega\sigma_k^2 m_k - (j\omega\sigma_k^2)^2}{2\sigma_k^2}} dx_k$$

$$= e^{j\omega m_k} e^{-\omega^2\sigma_k^2/2}$$

$$\Phi_{S_n}(\omega) = \prod_{k=1}^{n} e^{j\omega m_k} e^{-\omega^2\sigma_k^2/2} = e^{j\omega(m_1 + m_2 + \cdots + m_n) - \omega^2(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)/2}$$

$\Rightarrow$ Characteristic function of a Gaussian RV

Thus, S_n is a Gaussian RV with

Mean: $E[S_n] = m_1 + m_2 + \cdots + m_n$

Variance: $\mathrm{Var}[S_n] = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$

Example

- Find the pdf of a sum of n independent exponential RVs
- All exponential variables have parameter λ
- The characteristic function of a single exponential RV is

$$\Phi_X(\omega) = \frac{\lambda}{\lambda - j\omega}$$

- The characteristic function of the sum is

$$\Phi_S(\omega) = \left(\frac{\lambda}{\lambda - j\omega}\right)^n$$

- We can see that S_n is an m-Erlang RV

m - Erlang RV: $\alpha = m$, a positive integer

$$f_X(x) = \frac{\lambda(\lambda x)^{m-1} e^{-\lambda x}}{(m-1)!}, \qquad x > 0$$

→ Adding m independent exponentially distributed RVs with parameter λ

$$\Phi_X(\omega) = \left(\frac{\lambda}{\lambda - j\omega}\right)^m$$

Probability Generating Function (Recall)

- PGF of a nonnegative integer-valued RV N is defined by

$$G_N(z) = E[z^N] = \sum_{k=0}^{\infty} p_N(k) z^k$$

z – transform of PMF

- PMF of N is given by

$$p_N(k) = \frac{1}{k!} \frac{d^k}{dz^k} G_N(z) \bigg|_{z=0}$$

- Taking the first two derivatives of $G_N(z)$ and evaluating the result at $z = 1$, then we can find the first two moments of N

$$\frac{d}{dz} G_N(z)\bigg|_{z=1} = \sum_{k=0}^{\infty} p_N(k) k z^{k-1} \Bigg|_{z=1} = \sum_{k=0}^{\infty} k p_N(k) = E[N]$$

$$\frac{d^2}{dz^2} G_N(z)\bigg|_{z=1} = \sum_{k=0}^{\infty} p_N(k) k(k-1) z^{k-2} \Bigg|_{z=1} = \sum_{k=0}^{\infty} k(k-1) p_N(k) = E[N(N-1)] = E[N^2] - E[N]$$

$$E[N] = G_N'(1) \qquad VAR[N] = G_N''(1) + G_N'(1) - (G_N'(1))^2$$

Generating function of sum of discrete RVs

- Sum of n independent RVs

$$N = X_1 + X_2 + \cdots + X_n$$

- Generating function of N

$$G_N(z) = E[z^N] = E[z^{X_1+X_2+\cdots+X_n}]$$
$$= E[z^{X_1}] \cdots E[z^{X_n}]$$
$$= G_{X_1}(z) \cdots \Phi_{X_n}(z)$$

- Generating function of sum of n i.i.d random variables

$$G_N(z) = E[z^N] = E[z^{X_1+X_2+\cdots+X_n}]$$
$$= E[z^{X_1}] \cdots E[z^{X_n}]$$
$$= G_{X_1}(z) \cdots G_{X_n}(z)$$
$$= \{G_X(z)\}^n$$

Example

- Find the generating function for a sum of n independent, identically geometrically distributed RVs

- The generating function for a single geometric RV is given by

$$G_X(z) = \frac{pz}{1 - qz}$$

$$G_N(z) = \sum_{k=1}^{\infty} pq^{k-1} z^k = pz \sum_{k=1}^{\infty} (qz)^{k-1} = \frac{pz}{1 - qz}$$

- The generating function for a sum of n such independent RVs is

$$G_X(z) = \left\{\frac{pz}{1 - qz}\right\}^n$$

Good estimator 조건
1. $E[M_n] = \mu$
2. $E[(M_n - \mu)^2]$ is small (mean square error)
→ true mean에서 많이 벗어나지 말아야 함

Sample Mean

- Let X be an RV with $E[X] = \mu$ (unknown)

X의 분포를 모름 → 평균을 모르니
최대한 주사위 많이 던져서 평균을 취함

- $X_1, X_2, \ldots, X_n$ denote n independent, repeated measurements of X
- X_j's are independent, identically distributed (i.i.d.) RVs
- The sample mean of the sequence is used to estimate $E[X]$:

$$M_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

M_n is a function of n RVs
→ It is also RV

- μ is called true mean or ensemble average
- M_n is called the sample mean or empirical average (시간개념으로 표현)

empirical average
ensemble average

- Two statistical quantities for characterizing the sample mean's properties:
  - $E[M_n]$: we say M_n is unbiased if $E[M_n] = \mu$
  - $\mathrm{var}[M_n]$: we examine this value when n is large

Sample Mean

- The sample mean is an unbiased estimator for true mean μ

$$E[M_n] = E\left[\frac{1}{n}\sum_{j=1}^{n} X_j\right] = \frac{1}{n}\sum_{j=1}^{n} E[X_j] = \mu$$

- Since X_j's are i.i.d., the variance of $M_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ is

$$\mathrm{VAR}[M_n] = \frac{1}{n^2}\sum_{j=1}^{n} \mathrm{VAR}[X_j] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$\mathrm{VAR}[X] = \sigma^2$
(true variance)

as $n \to \infty$, $\mathrm{VAR}[M_n] \to 0$

→ The variance of the sample mean approaches 0

as the number of samples increases

→ This leads to the law of large numbers

law of large numbers is a theorem that describes the result of performing the same experiment a large number of times.

Weak Law of Large Numbers

- Let $X_1, X_2, \dots, X_n$ be a sequence of iid RVs with finite mean $E[X] = \mu$ and variance σ^2

- For any $\epsilon > 0$,

$$\lim_{n\to\infty} P[|M_n - \mu| < \varepsilon] = 1$$

→ For large enough n,

the sample mean will be close to the true mean with high probability

Proof) Apply Chebyshev inequality: $P(|X - \mu| \geq a) \leq \dfrac{\sigma_X^2}{a^2}$

0 as $n \to \infty$

$$P[|M_n - \mu| \geq \varepsilon] \leq \frac{\sigma^2/n}{\epsilon^2} \Rightarrow P[|M_n - \mu| < \varepsilon] \geq 1 - \cancel{\frac{\sigma^2}{n\varepsilon^2}}$$

Weak Law of Large Numbers

$$\lim_{n \to \infty} P[|M_n - \mu| < \varepsilon] = 1$$

- Interpretation:
  - ✓ Get n samples of X and the sample mean M_n.
  - ✓ The probability that $M_n - \varepsilon < \mu < M_n + \varepsilon$ is at least $1 - \frac{\sigma^2}{n\varepsilon^2}$

    true mean이 이 범위 안에 있을 확률
  - ✓ This interval can be used as a confidential interval (신뢰구간)
  - ✓ However, we still need to know σ^2, and the Chebyshev bound is not sufficiently tight.

Example

$n \to \infty$ 로 sample mean 취함 → true mean인 p 찾을 수 있음

- Toss a coin. We want to estimate the probability of a head p.
- Set $\epsilon = 0.05p$ with 95% confidence
- Howe many tossing a coin?

[Solution]

찾고 싶은건, 몇번을 동전을 던져야 sample mean 이 95% 신뢰도로 ϵ=0.05p 오차안으로 들어올것이냐?

$$\mathrm{P}[|M_n - p| < 0.05p] \to 0.95$$

- For a Bernoulli RV $\sigma^2 = p - p^2$
- $\frac{\sigma^2}{n\varepsilon^2} = \frac{p-p^2}{n(0.05p)^2} = \frac{1}{n(0.05)^2}\left(\frac{1}{p} - 1\right)$ must be equal to 0.05
- 실제 앞면 나올 확률이 $p = 0.5$ (a fair coin) 이었다고 가정해보자, $n = \frac{1}{(0.05)^3} = 20^3 = 8000$

This bound says we need at least 8000 samples for 95% confidence! Too many!

즉, 8000번 던지면 sample mean이 실제 p 값에 0.05p 오차 범위 안에 95% 신뢰도로 근접함... $|M_n - p| < 0.05p$

Example

- Consider a memoryless binary symmetric channel {0,1} with bit error rate p.
- Repeat each source bit N times, where N is an odd integer.
- Send 00...0 for 0, and Send 11...1 for 1.
- $P[E] = P[\text{more than} \frac{N}{2} \text{ errors}]$
  - For example, if we repeat 5 times. 3 or more errors will cause a decoding error at the receiver.

Example

$$X_1 + X_2 + \cdots + X_n > \frac{n}{2}$$

[Solution]

- Define

$$X_n = \begin{cases} 1 & \text{if the } n - \text{th transmission suffers an error} \\ 0 & \text{if otherwise} \end{cases}$$

$$P[E] = P\left[more\ than \frac{N}{2}\ errors\right] = P\left[\frac{X_1 + X_2 + \cdots + X_n}{n} > \frac{1}{2}\right] = P\left[M_n > \frac{1}{2}\right]$$

$$= P\left[M_n - p > \frac{1}{2} - p\right] \le \frac{VAR[M_n]}{\left(\frac{1}{2} - p\right)^2}$$

Bernoulli RV 분산=p(1-p)

$$\mathrm{VAR}[M_n] = \frac{VAR[X_n]}{n} = \frac{p(1-p)}{n} = \frac{pq}{n}$$

- For $p = 0.1$ and $N = 13$, we have the upperbound $P[E] \le 0.043$

Not too tight!

Strong Law of Large Numbers

- Let $X_1, X_2, \ldots, X_n$ be a sequence of iid RVs with finite mean $E[X] = \mu$ and finite variance, then

$$P\left[\lim_{n\to\infty} M_n = \mu\right] = 1$$

  - ✓ M_k is the sequence of sample mean computed using X_1 through X_k
  - ✓ With probability 1, every sequence of sample mean calculations will eventually approach and stay close to $E[X] = \mu$
  - ✓ The strong law implies the weak law

Central Limit Theorem

- Let $X_1, X_2, \dots, X_n$ be a sequence of arbitrary i.i.d RVs with finite mean $E[X] = \mu$ and finite variance σ^2
- Let S_n be the sum of the first n RVs in the sequences:

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

and define $$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Then $$\lim_{n\to\infty} Z_n \sim \mathcal{N}(0,1) \ \rightarrow \lim_{n\to\infty} P(Z_n \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx$$

- Equivalently,

$$\lim_{n\to\infty} S_n \sim \mathcal{N}(n\mu, n\sigma^2)$$

As n becomes large, S_n has Gaussian distribution!

결국, pdf의 convolution들이 Gaussian pdf로 converge함!

Proof of Central Limit Theorem

- First note that

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_k - \mu)$$

- The characteristic function of Z_n is given by

$$\Phi_{Z_n} = E\left[e^{j\omega Z_n}\right] = E\left[\exp\left(\frac{j\omega}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_k - \mu)\right)\right]$$

$$= E\left[\prod_{k=1}^{n} e^{j\omega(X_k-\mu)/\sigma\sqrt{n}}\right] \qquad e^{A+B+\cdots} = e^A e^B \ldots$$

$$= \left(E\left[e^{j\omega(X_k-\mu)/\sigma\sqrt{n}}\right]\right)^n \qquad (X_k\text{'s are iid})$$

Proof of Central Limit Theorem

- Expanding the exponential expression gives

$$E\left[e^{j\omega(X-\mu)/\sigma\sqrt{n}}\right] = E\left[1 + \frac{j\omega}{\sigma\sqrt{n}}(X-\mu) + \frac{(j\omega)^2}{2!\,n\sigma^2}(X-\mu)^2 + \cdots\right]$$

$$\approx 1 - \frac{\omega^2}{2n}$$

$$E\left[\frac{j\omega}{\sigma\sqrt{n}}(X-\mu)\right] = 0$$

$$E\left[\frac{(j\omega)^2}{2!\,n\sigma^2}(X-\mu)^2\right] = \frac{(j\omega)^2}{2!\,n\sigma^2}E[(X-\mu)^2] = \frac{(j\omega)^2}{2!\,n\sigma^2}\sigma^2$$

(the higher order term can be neglected as n becomes large)

Then we obtain

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

$$\Phi_{Z_n}(\omega) \to \left(1 - \frac{\omega^2}{2n}\right)^n \to e^{-\omega^2/2}, \quad \text{as } n \to \infty$$

$$\Phi_{X_k}(\omega) = e^{j\omega m_k} e^{-\omega^2\sigma_k^2/2}$$

$\mathcal{N}(0,1)$의
characteristic function
$= e^{-\omega^2/2}$

Example

- Mean: $E[X] = 1/\lambda$
- Variance: $\mathrm{var}[X] = 1/\lambda^2$

- The time between events is i.i.d exponential RVs with mean m sec
- Find the probability that the 1000th event occurs in time interval (1000±50)m
  - X_j is the time between events
  - S_n is the time of the n-th event (then $S_n = X_1 + X_2 + \cdots + X_n$)
  - $E[S_n] = nm$ and $\mathrm{VAR}(S_n) = nm^2$
- The CLT gives

$$P(950m \le S_{1000} \le 1050m) = P\left(\frac{950m - 1000m}{m\sqrt{1000}} \le Z_n \le \frac{1050m - 1000m}{m\sqrt{1000}}\right)$$

$$\approx \Phi(1.58) - \Phi(-1.58)$$

Example

- Packets arrive in a Poisson process with 10 packets per second.
- Find the probability that 650 or more packets arrive in a minute.

[Solution]

- $\lambda = 10 \text{ packets/sec} \rightarrow \alpha = \lambda t = 10 \times 60 = 600$
- $P[N = k] = \frac{600^k e^{-600}}{k!}$
- $P[N \geq 650] = 1 - P[N \leq 649] = 1 - \sum_{k=0}^{649} \frac{600^k e^{-600}}{k!}$

$\rightarrow$ Hard to compute!

Approach 1

- Let X be a Poisson RV which indicates $\#$ of arrivals during a second when the mean arrival rate is 10 packets per second. $\alpha = 10$
- Define X_j as the $\#$ of packets arriving during the j-th second.
- X_j's are independent and identical to X with $E[X] = 10$ and $VAR[X] = 10$
- Define $Z = X_1 + X_2 + \cdots + X_{60}$ as the $\#$ of packets arriving during a minute
- The probability that 650 or more packets arrive in a minute can be expressed as $P[Z \geq 650]$
- Take the approximation that Z is a Gaussian RV with mean $60E[X] = 600$ and variance $60VAR[X] = 600$

$$P[Z \geq 650] = P\left[\frac{Z - 600}{\sqrt{600}} \geq \frac{50}{\sqrt{600}}\right] \approx Q\left(\frac{50}{\sqrt{600}}\right) = 0.02061$$

Approach 2

Packet 하나 들어올때까지 걸리는 시간

10 packets/sec
→ 0.1 sec/1 packet

- Let X be a Exponential RV with mean 0.1 second.
- Define X_j as inter-arrival time of j-th packet.
- Then, X_j's are independent and identical to X
- We know $\mathrm{E[X]} = \frac{1}{\lambda} = \frac{1}{10}$, $\mathrm{VAR[X]} = \frac{1}{\lambda^2} = \frac{1}{100}$

적어도 650개가
60초안에는 발생함
을 의미

- Define $Z = X_1 + X_2 + \cdots + X_{650}$ (650개 event의 시간의 sum)
- Then, the probability that 650 or more packets arrive in a minute is $P[Z \leq 60]$
- Take the approximation that Z is a Gaussian RV with mean $650E[X] = 65$ and variance $650VAR[X] = 6.5$

$$P[Z \leq 60] = P\left[\frac{Z-65}{\sqrt{6.5}} \leq \frac{-5}{\sqrt{6.5}}\right] \approx Q\left(\frac{50}{\sqrt{650}}\right) = 0.02493$$

Two approaches yield values very close to each other.
Which is closer to the true value?

RV이 더 많이 더해졌으니
Gaussian 분포를 더 잘맞아떨어질것임

Homework

HW #1 - Weak Law of Large Numbers

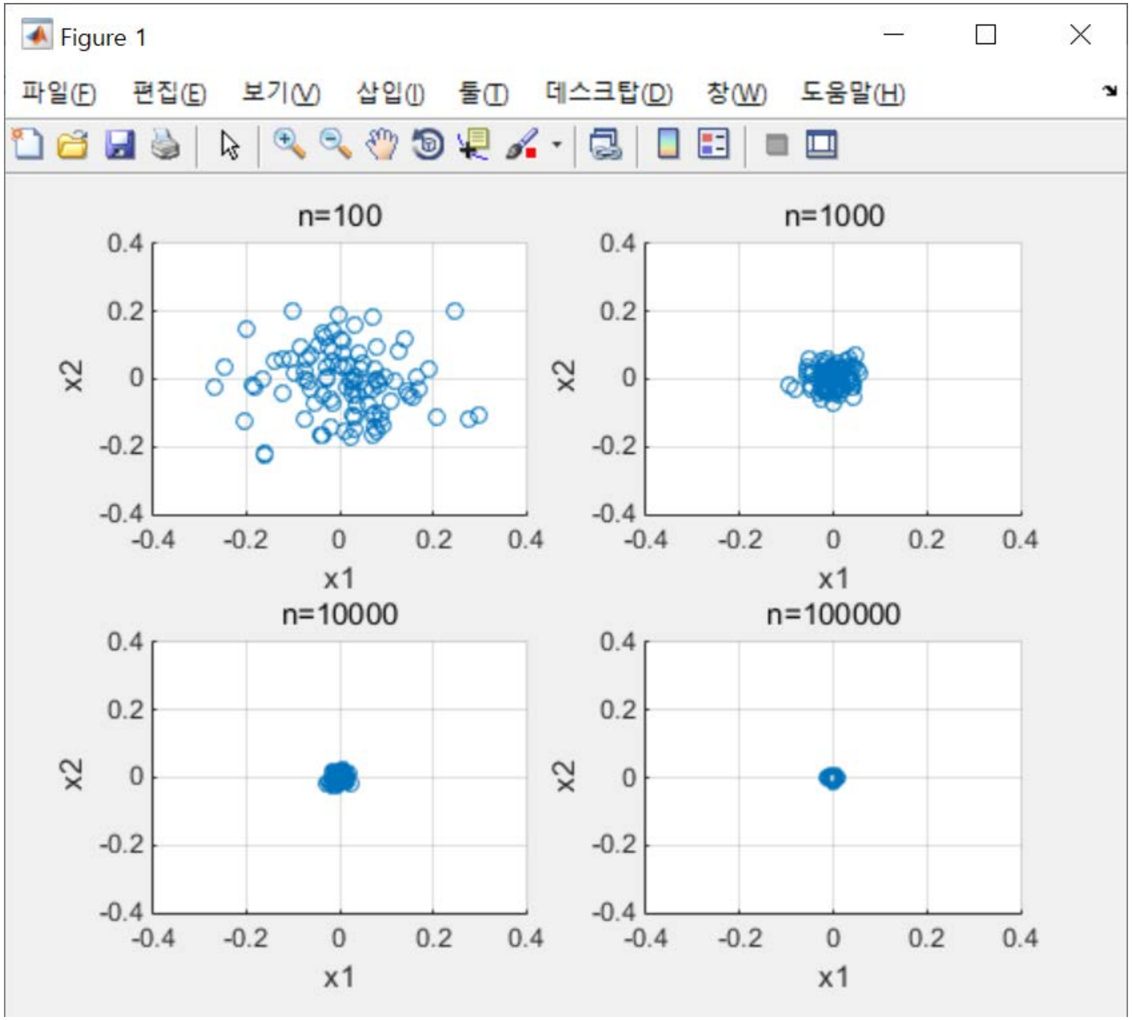
- 평균이 0, 분산이 1인 independent Gaussian RV X_1, X_2의 Joint PDF는 다음과 같다.

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi} e^{-\frac{(x_1)^2+(x_2)^2}{2}}, \qquad -\infty < x, y < \infty$$

- 다음 각 n sample의 sample mean의 scattergram을 그리시오.
  - i) $n = 100$
  - ii) $n = 1{,}000$
  - iii) $n = 10{,}000$
  - iv) $n = 100{,}000$
  - x축, y축의 범위는 각각 -0.4~0.4 까지만 display하시오.
  - scatter()함수를 사용하여 그림을 그릴것

Tip) randn() 함수를 사용하여 X_1, X_2를 각각 n개씩 동시에 생성한 다음 각기 평균을 내어, 2D 평면상에 점을 하나 찍음 → 이를 각기 100번 반복

HW #1 - Weak Law of Large Numbers

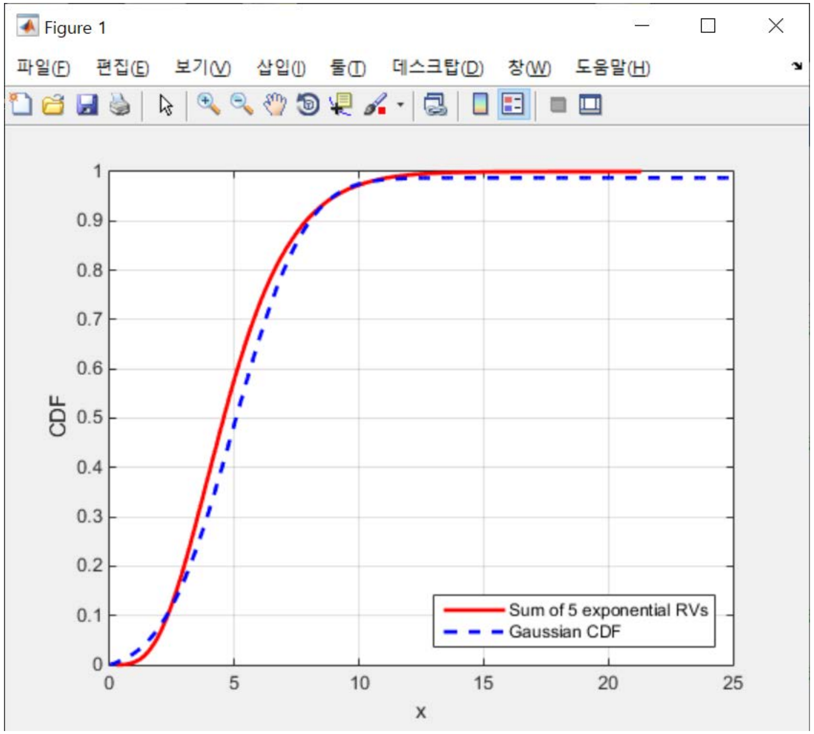


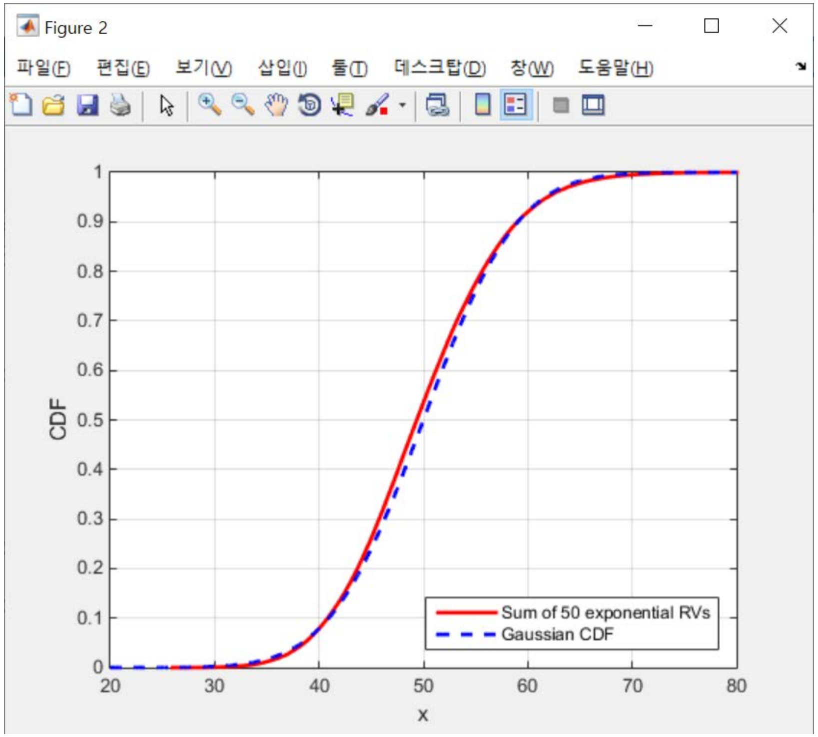
- As n increases, the probability of M_n's are concentrated at zero is high

HW #2 - CLT

- i.i.d exponential RVs의 n개 합에 대하여 Central Limit Theorem이 성립함을 확인하고자 한다. 다음을 수행하시오.
  - 그림1 ($n = 5$)
    - 평균이 $1/\lambda$인 exponential RV 5개 더하여 생성된 RV에 대한 empirical CDF를 그리시오.
    - 평균이 n/λ이고 분산이 n/λ^2인 Gaussian RV의 CDF를 그리시오.
    - $\lambda = 1$로 setting
    - x축은 $x = [0,25]$ 범위만 고려할 것
  - 그림2 ($n = 50$)
    - 평균이 $1/\lambda$인 exponential RV 50개 더하여 생성된 RV에 대한 empirical CDF를 그리시오.
    - 평균이 n/λ이고 분산이 n/λ^2인 Gaussian RV의 CDF를 그리시오.
    - $\lambda = 1$로 setting
    - x축은 $x = [20,80]$ 범위만 고려할 것

HW #2 - CLT



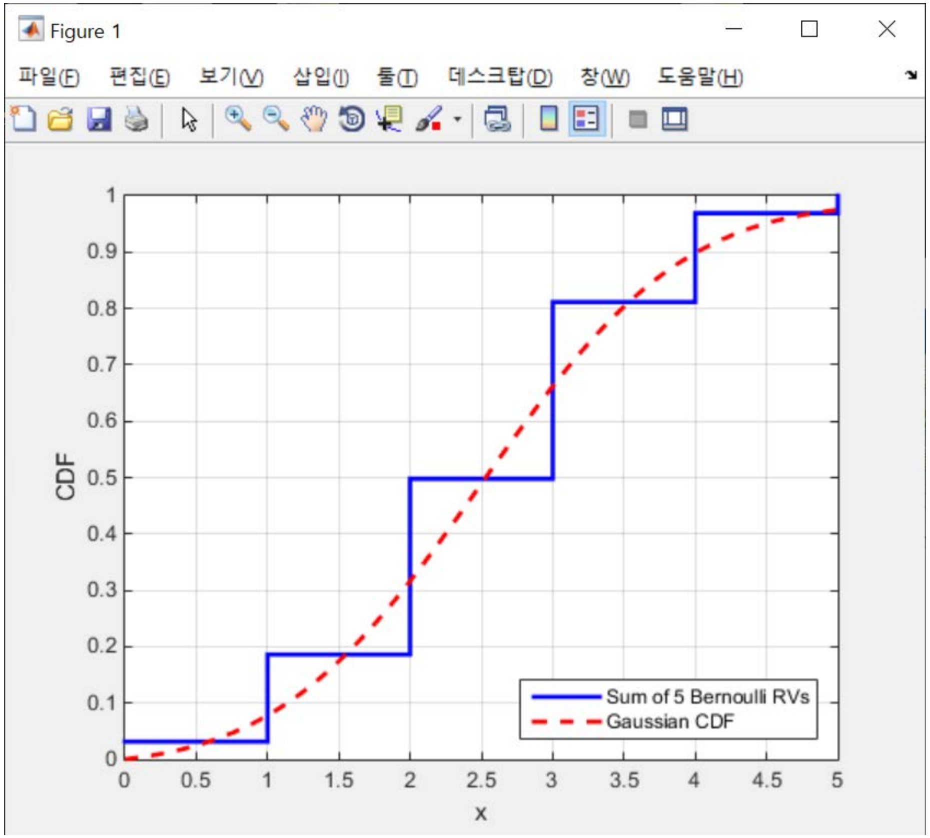


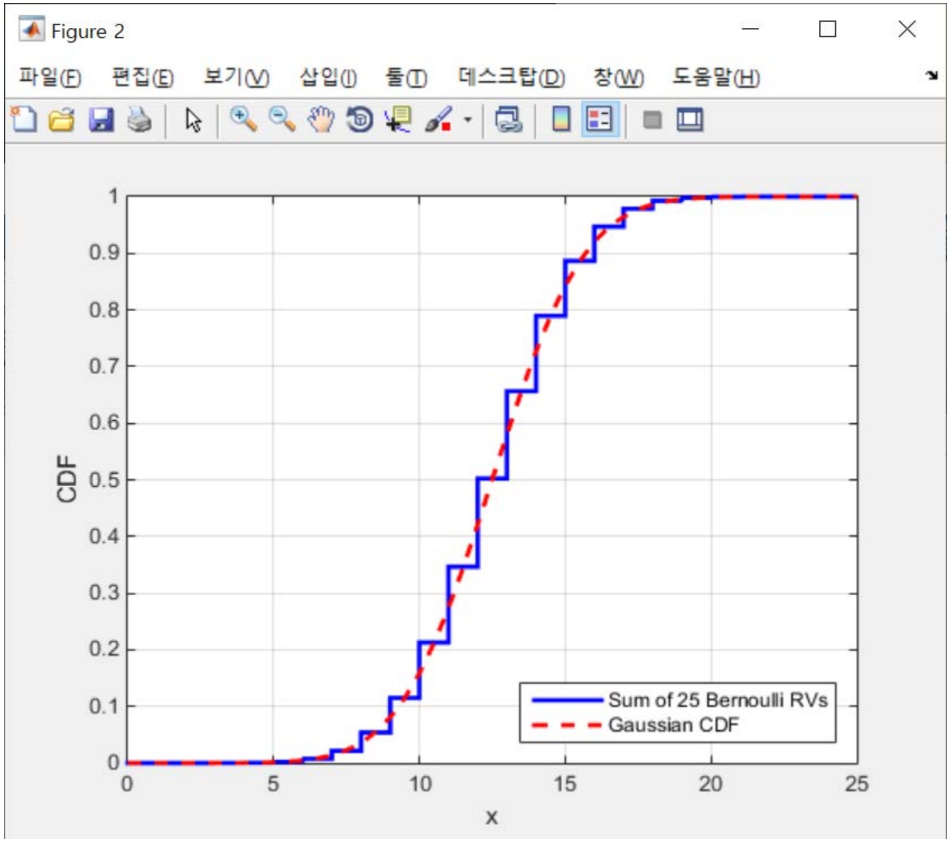
As n increases, the CDF approaches that of Gaussian distribution

HW #3 - CLT

- i.i.d Bernoulli RVs의 n개 합에 대하여 Central Limit Theorem이 성립함을 확인하고자 한다. 다음을 수행하시오.
  - 그림1 ($n = 5$)
    - $p = 1/2$인 Bernoulli RV 5개 더하여 생성된 RV에 대한 empirical CDF를 그리시오.
    - 평균이 np이고 분산이 $np(1-p)$인 Gaussian RV의 CDF를 그리시오.
    - x축은 $x = [0,5]$ 범위만 고려할 것
  - 그림2 ($n = 25$)
    - $p = 1/2$인 Bernoulli RV 25개 더하여 생성된 RV에 대한 empirical CDF를 그리시오.
    - 평균이 np이고 분산이 $np(1-p)$인 Gaussian RV의 CDF를 그리시오.
    - x축은 $x = [0,25]$ 범위만 고려할 것

HW #3 - CLT





As n increases, the CDF approaches that of Gaussian distribution

HW 제출

- 기한에 맞추어 e-class에 다음을 제출
  - Source code, 결과, 문제 풀이를 한글 또는 워드에 캡쳐 또는 붙여넣기 하여 하나의 파일로 제출할 것
  - 제출 파일명은 본인학번_이름.xxx 로 하여 제출할 것