Probability HW4 朱驛庭 111550093

Problem 1.

(a.)

(i) $\because X$ is a non-negative r.v. with $E[X] = \lambda T$, so by Markov's inequality,

we have $P(X > c\lambda T) \le \frac{E[X]}{c\lambda T} = \frac{\lambda T}{c\lambda T} = \frac{1}{c} \quad \forall c > 0$, $\therefore \underline{g_M(c;\lambda,T) = \frac{1}{c}}$

(ii) μ of $X = E[X] = \lambda T$, and $\sigma = \sqrt{Var[X]} = \sqrt{\lambda T}$, so by Chebyshev's inequality,

$P(|X - \lambda T| > k) \le \frac{\lambda T}{k^2}$. because we want to find $P(X > c\lambda T)$

and also $P(|X - \lambda T| > k) \ge P(X - \lambda T > k) = P(X > k + \lambda T)$.

let $k = (c-1)\lambda T$. so we have

$$P(X > c\lambda T) \le P(|X - \lambda T| < (c-1)\lambda T) \le \frac{\lambda T}{[(c-1)\lambda T]^2} = \frac{1}{(c-1)^2 \lambda T} \quad \#$$

$\therefore \underline{g_C(c;\lambda,T) = \frac{1}{(c-1)^2\lambda T}}$ #

(iii) MGF of X: $M_X(t) = e^{\lambda T(e^t - 1)}$, so by Chernoff bound, we have,

$P(X > c\lambda T) \le M_X(t) \cdot e^{-t(c\lambda T)} = e^{\lambda T(e^t-1) - t(c\lambda T)}$, let $\lambda T(e^t - 1) - t(c\lambda T) = g(t)$

$\frac{d}{dt} g(t) = \lambda T e^t - c\lambda T = 0$, $t = \ln c$, so minimum of $g(t) = \lambda T(c-1) - \ln c(c\lambda T)$

$= c\lambda T(1 - \ln c - \frac{1}{c})$

$\therefore \underline{g_F(c;\lambda,T) = e^{c\lambda T(1 - \ln c - \frac{1}{c})}}$ #

(c.) for each $i = 1, 2, \dots, N$, define X_i to be Bernoulli r.v. for which $X_i = 1$ when it return the correct answer at the i-th trial. So $X_i \sim$ Bernoulli$(\frac{1}{2} + \delta, \frac{1}{2} - \delta)$

Define $\bar{X} = \frac{1}{N}(X_1 + X_2 + \dots + X_N)$, then by negative part of Hoeffding's inequality.

Choose $\varepsilon = \delta$. $P(\bar{X} - (\frac{1}{2} + \delta) < -\delta) \le \exp(-2N\delta^2) \Rightarrow P(\bar{X} < \frac{1}{2}) \le \exp(-2N\delta^2)$

so if $N \ge \frac{1}{2\delta^2}\ln(\frac{1}{\varepsilon})$. $P(\bar{X} < \frac{1}{2}) \le \exp(-2N\delta^2) \le \exp(-2\delta^2 \cdot \frac{1}{2\delta^2}\ln(\frac{1}{\varepsilon})) = \varepsilon$

So the answer is wrong with probability at most ε when $N \ge \frac{1}{2\delta^2}\ln(\frac{1}{\varepsilon})$

In other words, the answer is correct with probability at least $1 - \varepsilon$,

when $N \ge \frac{1}{2\delta^2}\ln(\frac{1}{\varepsilon})$

Problem 2.

(a) • Derive the marginal PDF of Z_1 by taking the integration of $f_{Z_1,Z_2}(z_1,z_2)$ over z_2.

let's break down $f_{X_1X_2}(x_1,x_2)$ first, we can separate $f_{X_1X_2}(x_1,x_2)$ into two parts. The first part is nothing related to z_2, so it can be seen as constant when we taking integration over z_2. And the second part is the remainings. So we have:

$$f_{z_1z_2}(z_1,z_2) = \underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]\right)}_{\text{(part 1)}}\underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right]\right)}_{\text{(part 2)}}$$

by the definition of marginal PDF:

$$f_{z_1}(z_1) = \int_{-\infty}^{+\infty} f_{z_1z_2}(z_1,z_2)\,dz_2$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]\right)\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right]dz_2$$

let $\sigma = \sigma_2\sqrt{1-\rho^2}$. $\mu = \mu_2 + \rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)$ $\quad \rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1-\mu_2)$ $-z_2$

then $\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right]$ can be written as

$\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(z_2-\mu)^2}{2\sigma^2}\right]$. This is normal distribution

$$\Rightarrow \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right] \sim N\left(\mu_2+\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1),\sigma_2^2(1-\rho^2)\right)$$

$$\text{so } f_{z_1}(z_1) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]\cdot(1) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]$$

then we know that marginal PDF of z_1 is also normal

- Show that conditioned on that $Z_1 = z_1$, the conditional distribution of Z_2 is normal with mean $\mu_2 + \frac{\rho\sigma_2(z_1-\mu_1)}{\sigma_1}$ and variance $(1-\rho^2)\sigma_2^2$. (Hint: Follow the definition of conditional PDF)

taking the result from previous problem, we already know that

$$f_{z_1 z_2}(z_1, z_2) = \left(\frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]\right)\left(\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right]\right)$$

and $f_{z_1}(z_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(z_1-\mu_1)^2}{2}\right]$

by the definition of conditional PDF :

$$f_{z_2|z_2}(z_2|z_1) = \frac{f_{z_1 z_2}(z_1, z_2)}{f_{z_1}(z_1)}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma_2^2}\left(\frac{\rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1)-(z_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2\right]$$

$$\sim N\left(\mu_2 + \rho\cdot\frac{\sigma_2}{\sigma_1}(z_1-\mu_1),\ \sigma_2^2(1-\rho^2)\right) \quad \#$$

Problem 3.

we already know X_n converges to a in probability and Y_n converges to b in probability. So $\exists N_1, N_2 > 0$. $\forall \varepsilon' > 0$ s.t. if $n \geq N_1 \Rightarrow P(\{\omega : |X_n(\omega)-a| \geq \varepsilon'\}) = 0$ if $n > N_2 \Rightarrow P(\{\omega : |Y_n(\omega)-b| \geq \varepsilon'\}) = 0$. let $N = \max\{N_1, N_2\}$, $\varepsilon' = \varepsilon$, then

if $n > N$, $P(\{\omega : |X_n(\omega)+Y_n(\omega)-(a+b)| \geq \varepsilon\}) \overset{\text{triangle inequality}}{\leq} P(\{\omega : |X_n(\omega)-a| + |Y_n(\omega)-b| \geq \varepsilon\})$

$\underset{\text{union bound}}{\leq} P(\{\omega : |X_n(\omega)-a| \geq \varepsilon\}) + P(\{\omega : |Y_n(\omega)-b| \geq \varepsilon\}) = 0.$

so $\{X_n + Y_n\}_{n=1}^{\infty}$ converges to $(a+b)$ in probability.

Problem 4.

(a) consider V_n. if dogecoin rises on that day $V_{n+1} = \underbrace{(0.5 \times 1.5 + 0.5) V_n}_{1.25 V_n}$

if it drops, $V_{n+1} = (0.5 \times 0.7 + 0.5) = 0.85\, V_n$. so :

$\frac{V_{n+1}}{V_n} = \begin{cases} 1.25 & \text{w.p.} = \frac{1}{2} \\ 0.85 & \text{w.p.} = \frac{1}{2} \end{cases}$, let $R_i := \frac{V_{i+1}}{V_i}$, so $E[R_i] = 1.25 \times 0.5 + 0.85 \times 0.5$
$= 1.05$

so by SLLN, $P(\{\omega : \lim_{n\to\infty} \frac{R_n(\omega)}{n} = 1.05\}) = 1$. in other word, the total fortune will continue to grow at a rate of 5% in average when n is big enough So $V_n \to \infty$ as $n \to \infty$.

(b.) Define the daily return factor as X_i, representing the change in the value of the investment each day. So :

$$X_i = \begin{cases} 1.5 & w.p. = \frac{1}{2} \\ 0.7 & w.p. = \frac{1}{2} \end{cases} \quad \text{and} \quad V_{n+1} = V_n \left((1-\alpha) + \alpha \cdot X_{n+1} \right)$$

$$\Rightarrow V_1 = V_0 \left((1-\alpha) + \alpha \cdot X_1 \right), \text{ so } V_n = V_0 \prod_{i=1}^{n} \left((1-\alpha) + \alpha \cdot X_i \right)$$

$$\log\left(\frac{V_n}{V_0}\right) = \log\left(\prod_{i=1}^{n} \left[(1-\alpha) + \alpha \cdot X_i \right] \right) = \log\left[(1-\alpha)+\alpha X_1\right] + \log\left[(1-\alpha)+\alpha X_2\right] + \cdots$$

$$+ \log\left[(1-\alpha) + \alpha X_n\right] = \sum_{i=1}^{n} \log\left[(1-\alpha)+\alpha X_i\right], \text{ define } R_i := \log\left[(1-\alpha)+\alpha X_i\right]$$

because X_i is i.i.d, so R_i is also i.i.d. denote $\sum_{i=1}^{n} R_i$ as S_n, and

$E[R_i] = \frac{1}{2}\log(1+0.5\alpha) + \frac{1}{2}\log(1-0.3\alpha)$, then by SLLN we know

$$P\left(\left\{ \omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \frac{1}{2}\log(1+0.5\alpha) + \frac{1}{2}\log(1-0.3\alpha) \right\}\right) = 1.$$

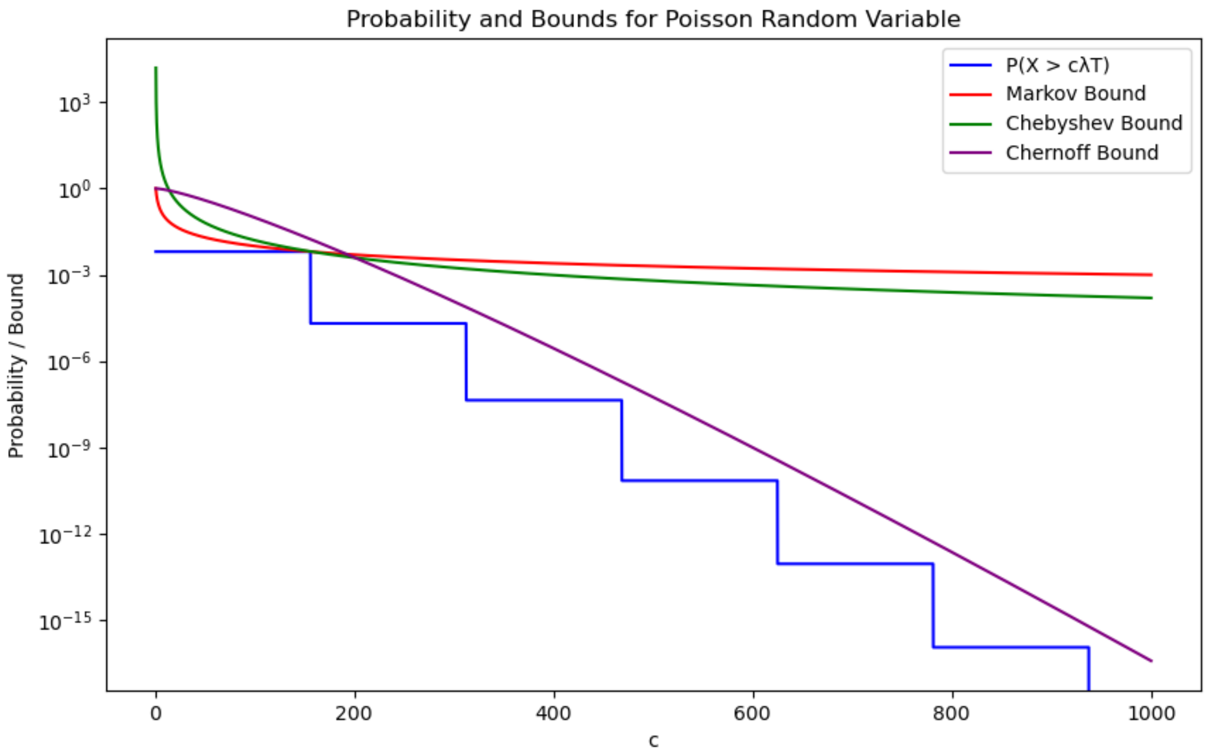
in other words. $\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\log\left[(1+0.5\alpha)(1-0.3\alpha)\right]$

also $\log V_n = \log V_0 + \sum_{i=1}^{n} R_i = 3 + S_n$, so

$$\therefore \text{ as } n \to \infty, \quad \frac{\log V_n}{n} = \frac{3+S_n}{n} \to \frac{1}{2}\log\left[(1+0.5\alpha)(1-0.3\alpha)\right]$$

Problem 1-b

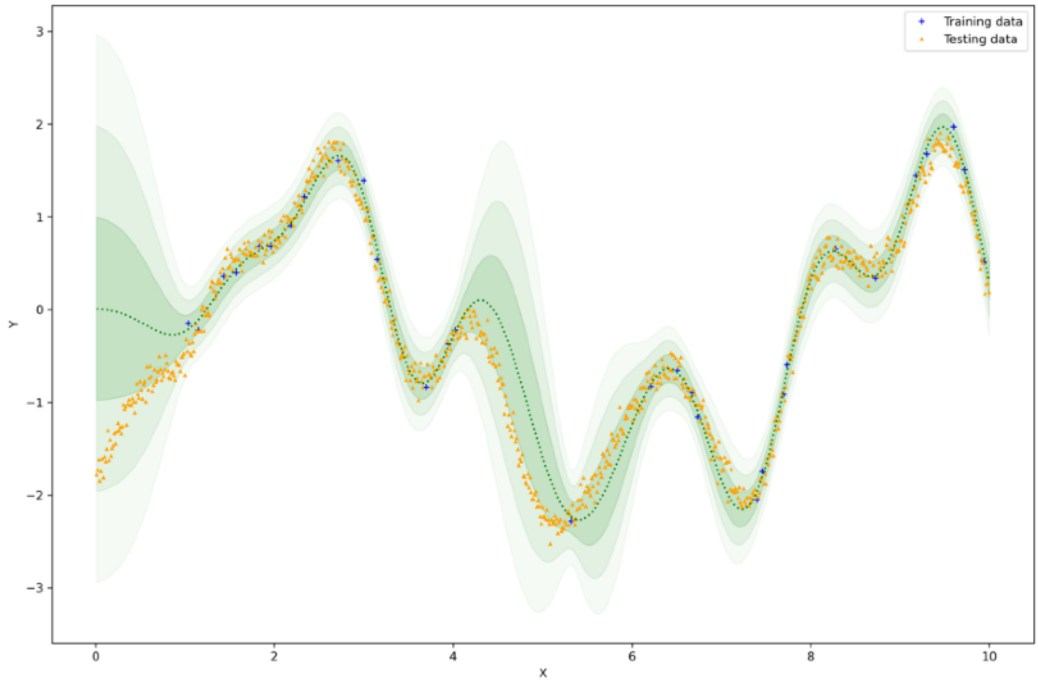
Result:



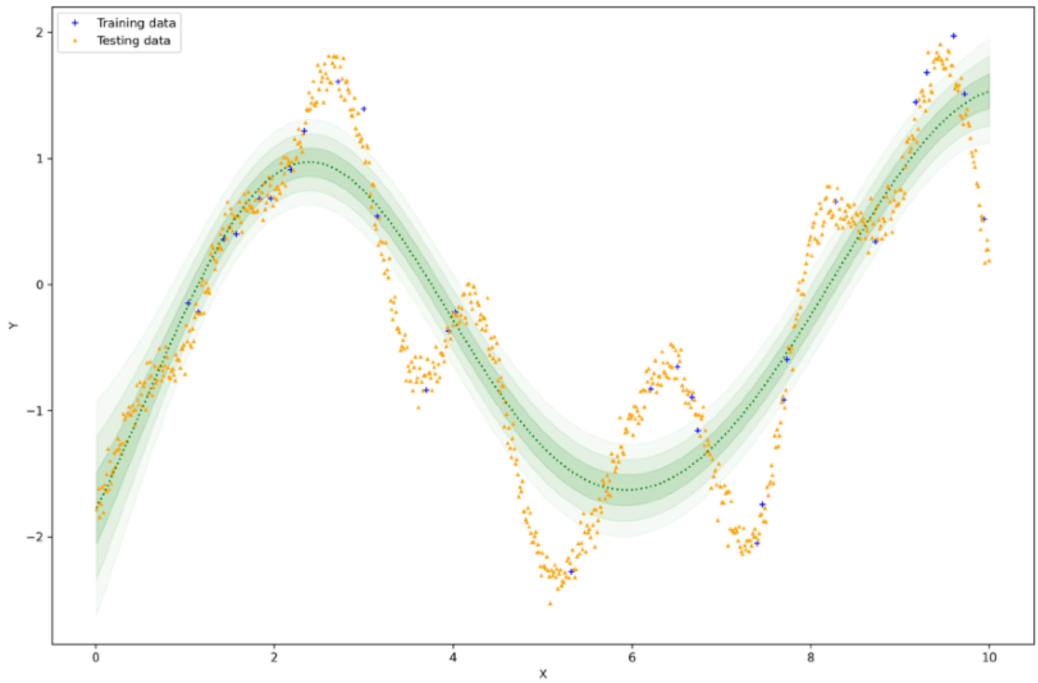
As we can see in the picture. Markov's bound (red line) is the least tight, meaning it's generally further away from the true probability. And Chebyshev's Bound (green line) is a little bit tighter than Markov's. Chernoff bound (purple line) is the tightest bound, which capture the decreasing rate of true probability.

Problem 2-b

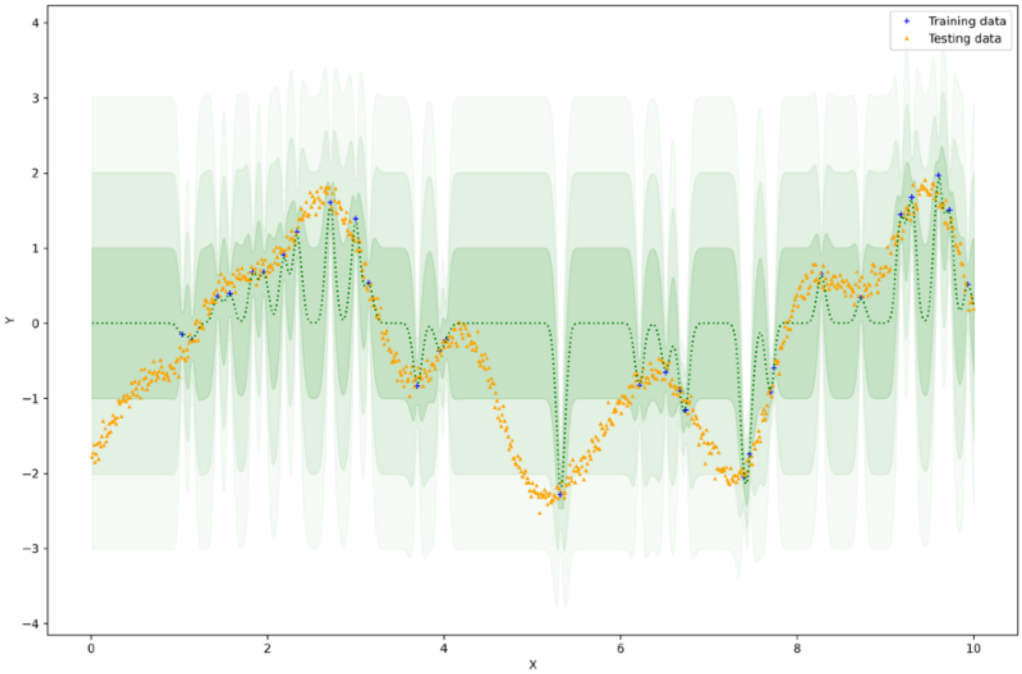
Result:



L = 0.5



L = 2.5

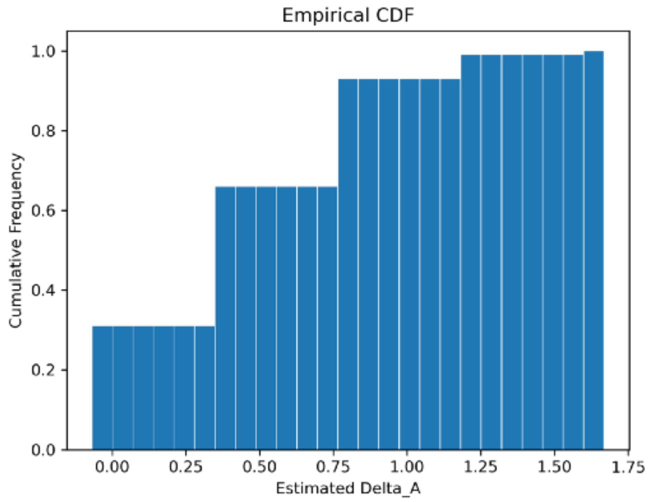


L = 0.05

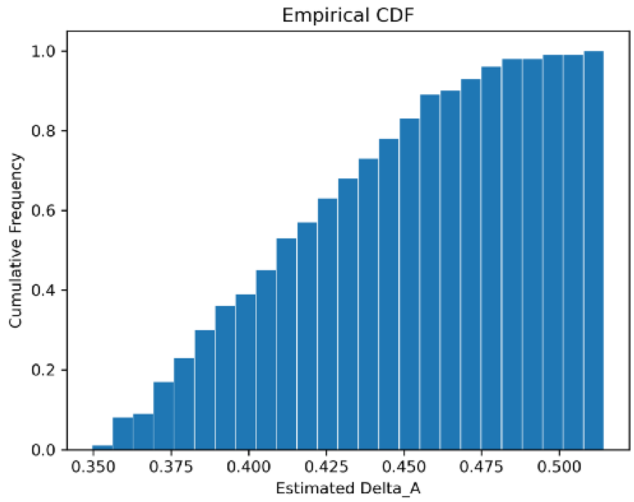
The prediction result of the testing dataset under $\sigma_f = 1$, $\sigma = 0.1$, $l = 0.5,\ 2.5,\ 0.05$ is showing below. When $l = 0.5$, it seems like the predicate function is the closest to the testing datas, as every datas are within 3 of standard deviation. And when $x \in [0,1]$, it seems most uncertain to me. When $l = 2.5$, it generally follow the increasing and decreasing rate of testing datas, but not much precise compared to $l = 0.5$. When $l = 0.05$, the standard deviation is too large so it don't even capture the increasing and decreasing rate of testing datas.

Problem 5

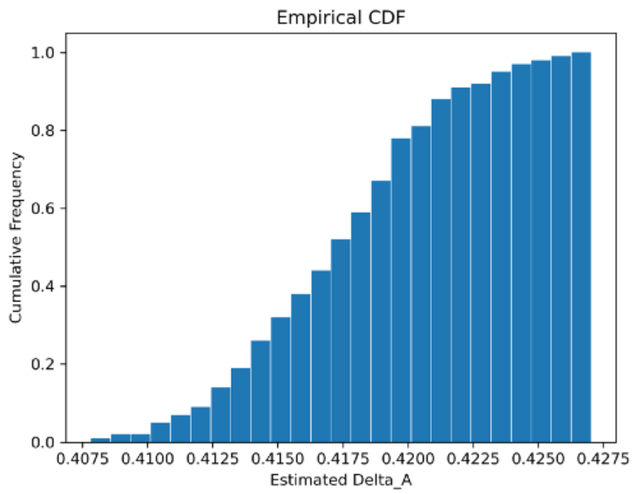
Result:



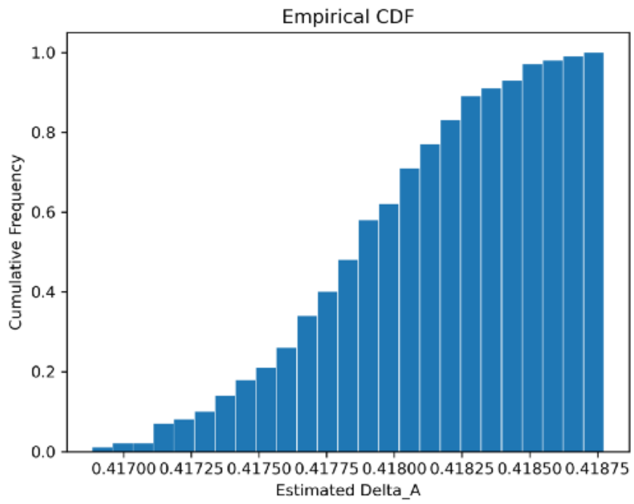
N=10^1



N=10^3



N=10^5



N=10^7

As we can see in the results, the possible values of delta_a under different N are showing below. When N = 10^1, the most possible value of delta_a is around 0.25 ~ 1.25. When N = 10^3, the most possible value of delta_a is around 0.35 ~ 0.475. When N = 10^5, the most possible value of delta_a is around 0.415 ~ 0.4225. When N = 10^7, the most possible value of delta_a is around 0.4175 ~ 0.41825.