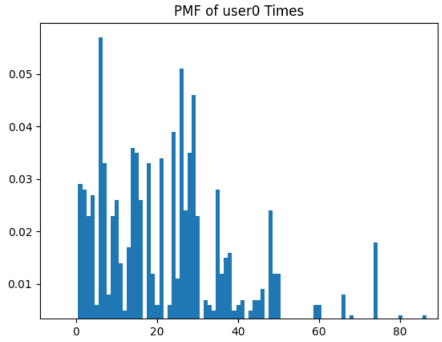
Q1(a)



Q1(b)

Mapping the values for X_0 to all the values for user0 we get 582 values where $X_0 = 1$. Since $Prob(X_0 = 1) = E[X_0]$. We can use the formula for empirical mean:

$$\frac{1}{N}\sum_{k=1}^{N} X_k = \frac{582}{1000}$$

$Prob(X_0 = 1) = 0.582$

Q1(d)

Code in appendix

Q1(c)

Chebyshev:

- Gives full distribution of X_0
- Only requires mean and variance to full describe distribution
- Con: Approximation when N is finite, hard to determine accuracy

$$\mu = 0.582, \sigma = \sqrt{\mu(1-\mu)} = 0.493, N = 1000$$

$$\mu - \frac{\sigma}{\sqrt{0.05N}} \leq X_0 \leq \mu + \frac{\sigma}{\sqrt{0.05N}} = 0.582 - \frac{0.493}{\sqrt{0.05(1000)}} \leq X_0 \leq 0.582 + \frac{0.493}{\sqrt{0.05(1000)}}$$

$$0.512 \leq X_0 \leq 0.651$$

CLT:

- Provides an actual bound and not an approximation
- Works for all N
- Con: It's loose in general

$$\frac{\frac{(X_1+X_2+...+X_n)}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} =$$

Bootstrapping:

- Gives full distribution without assuming normality
- Con: Approximation when N is finite, hard to determine accuracy
- Con: Requires the availability of all N measurements

Q2

user1: 0.416 | **user2**: 0.399 | **user3**: 0.334

Q3

Using marginalisation and summing all the probabilites to get Z_n:

$$P(X_0 = 1)P(U_0) + P(X_1 = 1)P(U_1) + P(X_2 = 1)P(U_2) + P(X_3 = 1)P(U_3)$$

$$0.582(0.09742...) + 0.416(0.40468...) + 0.399(0.23529...) + 0.334(0.26260...)$$

$Z_n = 0.4066392298682297$

Q4

$$P(U_n = 0|Z_n > 10ms) = P(E|F) = \frac{P(F|E)P(E)}{P(F)}$$

- $P(F|E) = 0.582$ (from Q1)
- $P(E) = 0.09742483650256$ (from top line of dataset)
- $P(E^c) = 1 - P(E) = 1 - 0.09742483650256 = 0.902575163$
- $P(F|E^c) = 1 - P(F|E) = 1 - 0.582 = 0.418$
- $P(F) = P(F|E) * P(E) + P(F|E^c) * P(E^c)$

$$P(U_n = 0|Z_n > 10ms) = \frac{0.582 * 0.09742483650256}{(0.582 * 0.09742483650256) + (0.418 * 0.902575163)}$$

$P(U_n = 0|Z_n > 10ms) = 0.1306547741392685$