① 1.

This question is regarding iid Bernoulli samples.
The likelihood function of such distribution is
$L(x_1,..,x_n;p) = p^{\Sigma(x_i)} \cdot (1-p)^{(n-\Sigma(x_i))}$ where n is the total number of samples and $\Sigma(x_i)$ is the sum of $x_1,..,x_n$.

The log-likelihood function gives
$\log L(x_1,..,x_n;p) = \log(p^{\Sigma(x_i)} \cdot (1-p)^{(n-\Sigma(x_i))})$

Using log properties, we can simplify it into
$$= \sum_{i=1}^{n}(x_i)\log p + \left(n - \sum_{i=1}^{n} x_i\right)\log(1-p)$$

To find the maximum likelihood estimator, we take the derivative with respect to p and set it equal to 0

$$\frac{d}{dp}\left(\sum_{i=1}^{n} x_i \log p + \left(n-\sum_{i=1}^{n} x_i\right)(\log(1-p))\right) = 0$$

$$\frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p} = 0$$

$$\left(\sum_{i=1}^{n} x_i\right)(1-p) = \left(n - \sum_{i=1}^{n} x_i\right)(p)$$

$$\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \cdot p = pn - \sum_{i=1}^{n} x_i \cdot p$$ → the $\sum_{i=1}^{n} x_i \cdot p$ cancels out

$$\sum_{i=1}^{n} x_i = pn$$

$$p_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Now, we have to prove that P_{MLE} is maximum, not minimum, which can be computed by doing a second derivative of the original function.

If the second derivative is negative, it means that the graph is concave up, meaning that it is maximum at that point.

$$\frac{d}{dp}\left(\frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p}\right)$$

$$= \frac{-\sum_{i=1}^{n} x_i}{p^2} - \frac{n-\sum_{i=1}^{n} x_i}{(1-p)^2} < 0 \rightarrow \text{(factor -1 from each)}$$

$$\frac{\sum_{i=1}^{n} x_i}{p^2} + \frac{n-\sum_{i=1}^{n} x_i}{(1-p)^2} > 0 \text{ (change the inequality)}$$

↳ Since p is between 0 and 1 by the definition of probability, all denominator value is positive.

↳ The sum of all x_i cannot be negative since each individual x_i is either 0 or 1

↳ n - sum of all x_i cannot be negative since $n \geq$ sum of all x_is.

Therefore, we get $\frac{\text{positive value}}{\text{positive value}} + \frac{\text{positive value}}{\text{positive value}}$, which is always greater than 0, meaning that the P_{MLE} is a maximum and not minimum.

① 2.

The joint probability function of Y and p

$p(Y=y, p) = p(Y=y|p) \cdot P(p)$

where $p(Y=y|p)$ is the likelihood function of Y given p and $P(p)$ is the prior distribution

Assuming that the prior p is beta(α, β), the PDF of

$p(p)$ is $\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$

$=$

It is given that X_i's are independent and have same Bernoulli (p) distribution. Therefore, the $P(Y=y|p)$

$= p(Y=y_1|p) \cdot p(Y=y_2|p) \cdot \ldots \cdot p(Y=y_n|p)$

$= {}_nC_y \cdot p^y \cdot (1-p)^{n-y}$

Combining two functions, we get

$P(Y=y, p) = {}_nC_y \cdot p^y \cdot (1-p)^{n-y} \cdot \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot p^{\alpha-1} \cdot (1-p)^{\beta-1}$

$= {}_nC_y \cdot p^{(y+\alpha-1)} \cdot (1-p)^{(n-y+\beta-1)} \cdot \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)}$

① 3.

This question is regarding normal distribution.
The likelihood of such distribution is

$$L(x_1,..,x_n ; N(\mu,\sigma^2)) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i-\mu)^2/2\sigma^2}\right)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2} = (2\pi\sigma^2)^{-n/2} \cdot e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2}$$

If we take the log-likelihood function,

$$\log(L(x_1,...,x_n ; N(\mu,\sigma^2)) = \frac{-n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2$$

If we take derivative with respect to μ and set it equal to 0, we get

$$2\sum_{i=1}^{n}(x_i-\mu)/2\sigma^2 = 0$$

$$\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\sum_{i=1}^{n} x_i - n\mu = 0$$

$$\boxed{\mu_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

If we take derivative with respect to σ^2 and set it equal to 0, we get

$$\frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4} = 0$$

$$\frac{-n\sigma^2}{2\sigma^4} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4} = 0$$

$$-n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\boxed{\sigma^2_{MLE} = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}}$$

① 4.

The first moment of Exponential (λ) is the mean, which is $\frac{1}{\lambda}$

The mean of the sample is $\frac{1}{n} \cdot \sum_{i=1}^{n} x_i$

Since the first sample moment is the sample mean,

$$\frac{1}{\lambda} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

$$\boxed{\hat{\lambda}_{MoM} = \frac{n}{\sum_{i=1}^{n} x_i}}$$

① 5(a)

This question is regarding $\beta(\theta, 1)$

The likelihood function for such distribution is

$$L(X_1,\dots,X_n; \beta(\theta,1)) = \prod_{i=1}^{n} \frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)} (x_i)^{\theta-1} \cdot (1-x_i)^{1-1}$$

$$= \prod_{i=1}^{n} \frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)} (x_i)^{\theta-1}$$

According to wikipedia, we can denote $\Gamma(n)$ as $(n-1)!$ if n is a real positive integer.

The piazza post reply mentioned that θ can be real positive here and since the sum of two real positive integers always gives a real positive integer, $\theta+1$ is also real positive.

$$\therefore \quad = \prod_{i=1}^{n} \frac{(\theta)!}{(\theta-1)!\,0!} (x_i)^{\theta-1}$$

$$= \prod_{i=1}^{n} \theta \cdot x_i^{\theta-1}$$

The log-likelihood function gives

$$\log(L(X_1,\dots,X_n; \beta(\theta,1))) = \log\left(\prod_{i=1}^{n} \theta \cdot x_i^{\theta-1}\right)$$

$$= \log(\theta^n) + \sum_{i=1}^{n} \log x_i^{\theta-1}$$

$$= n\log(\theta) + \sum_{i=1}^{n} \log x_i^{\theta-1}$$

$$= n\log(\theta) + (\theta-1)\sum_{i=1}^{n} \log(x_i)$$

Taking derivative with respect to θ, and setting it equal to 0 gives

$$\frac{n}{\theta} + \sum_{i=1}^{n} \log(x_i) = 0$$

$$\boxed{\theta_{MLE} = \frac{n}{-\sum_{i=1}^{n} \log(x_i)}}$$

→ since the value of all $\log(x_i)$s are between 0 and 1, they sum up to a negative value. multiplying the sum to (-1) gives a positive θ_{MLE}

① 5 (B)

The first moment of $B(\theta,1)$ is the mean, which is $\frac{\theta}{\theta+1}$ according to the hint.

The mean of the sample is $\frac{1}{n}\sum_{i=1}^{n} x_i$

Since the first sample moment is the sample mean,

$$\frac{\theta}{\theta+1} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\theta = (\theta+1)\left(\frac{1}{n}\right)\sum_{i=1}^{n} x_i$$

$$= \frac{\theta}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$-\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{\theta}{n}\sum_{i=1}^{n} x_i - \theta$$

$$= \theta\left(\frac{1}{n}\sum_{i=1}^{n} x_i - 1\right)$$

$$\boxed{\therefore\ \theta_{MOM} = \frac{\left(\frac{-1}{n}\right)\left(\sum_{i=1}^{n} x_i\right)}{\left(\frac{1}{n}\sum_{i=1}^{n} x_i - 1\right)}}$$

② Given that there are n people at the party, each person can shake hands with n-1 people (each person cannot shake hands by himself/herself).

Let X_i be the random variable that represents the number of handshakes for each person i.

Since all X_is follow a binomial distribution with $p = 1/10$, $E(X_i) = n-1/10$

The question asks to prove that every person tends to shake hands with probability of 1 in range $[0.95\,n/10, 1.05\,n/10]$ as $n \to \infty$.

As $n \to \infty$, the expected value is assymptotic to $n/10$ since $n-1 \approx n$.

The question gave the range as $[0.95\,n/10, 1.05\,n/10]$, which means that the range is $\frac{n}{10} \pm 0.05\,n/10$.

↳ therefore, $\delta = 0.05$ in this question.

This means that we can subtract the probability of a person that tends to not shake hands in range $[0.95\,n/10, 1.05\,n/10]$ as $n \to \infty$.

Let A_i be the random variable that person i does not tend to shake hands in range $[0.95\,n/10, 1.05\,n/10]$ as $n \to \infty$.

↳ A_i also follows the same binomial distribution as X_i, and has the same expected value of $n/10$ as $n \to \infty$.

Therefore, we are looking for $1 - P(A_1 \cup A_2 \cup \ldots \cup A_n)$

↱ Since all A_is follow the same distribution

According to the union bound, $P(A_1 \cup A_2 \cup \ldots \cup A_n) \leq nP(A_1)$

In other words, $P(A_1 \cup A_2 \cup \ldots \cup A_n)$ has the upper bound of $nP(A_1)$

Now, we have to prove that $n \cdot Pr(A_1)$ goes to 0 as $n \to \infty$, meaning that $Pr(A_1)$ goes to 0 faster than n.

↳ $Pr(A_1) \ll \frac{1}{n}$ as $n \to \infty$.

↳ If the upper bound $n \cdot Pr(A_1)$ goes to 0, $Pr(A_1 \cup A_2 \cup \ldots \cup A_n)$ will also go to 0 as $n \to \infty$.

We can now apply Chernoff bound on $P(A_1)$

$$Pr(|A_1 - n/10| \geq 0.05(n/10)) \leq 2e^{-(0.05)^2/3 \cdot n/10}$$
$$= 2e^{-n/12000}$$

According to Chernoff bound, $P(A_1)$ has the upper bound of $2e^{-n/12000}$ which goes to 0 as $n \to \infty$ faster than $\frac{1}{n}$

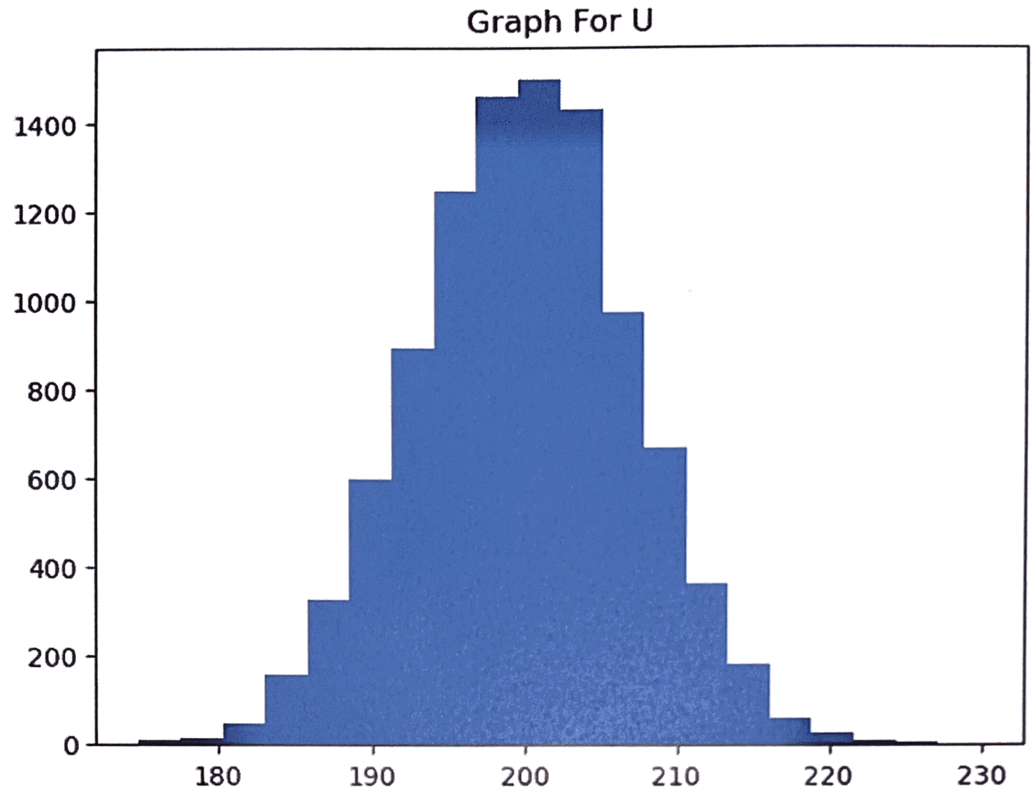
In other words, $\lim_{n \to \infty} \frac{2n}{e^{n/12000}} \to 0$ as $n \to \infty$ since $e^{n/12000}$ is exponential.

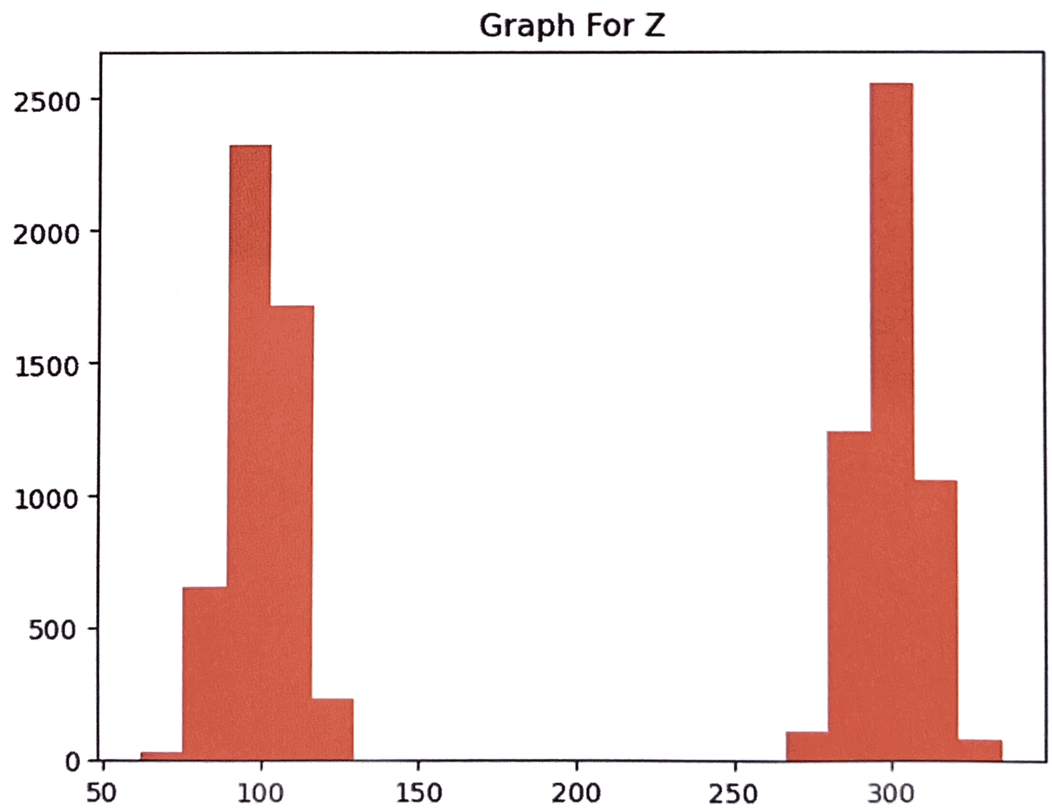
↳ this means that $P(A_1)$ also goes to 0 faster than $1/n$ as $n \to \infty$.

Therefore, $n P(A_1) = 0$ as $n \to \infty$, which means that $P(A_1 \cup A_2 \cup \ldots \cup A_n)$ also goes to 0 as $n \to \infty$, since the upper bound of it goes to 0.

$\therefore$ $1 - Pr(A_1 \cup A_2 \cup \ldots \cup A_n) = 1 - 0 = 1$ as $n \to \infty$, meaning that the probability that every person at the party shook hands in the range $[0.95\, n/10, 1.05\, n/10]$ tends to be 1 as $n \to \infty$.

③ 1.


Graph For U
0
200
400
600
800
1000
1200
1400
180
190
200
210
220
230


Graph For Z
0
500
1000
1500
2000
2500
50
100
150
200
250
300

③ 2.

Expected value of U

$U = \frac{1}{2}(X+Y)$

$E(U) = E(\frac{1}{2}(X+Y))$

$= E(\frac{1}{2}X + \frac{1}{2}Y)$

Since X and Y are independent,

$E(U) = E(\frac{1}{2}X) + E(\frac{1}{2}Y)$

$= \frac{1}{2}E(X) + \frac{1}{2}E(Y)$

The expected values of X and Y are their means.

$E(U) = \frac{1}{2}(100) + \frac{1}{2}(300)$

$= 200$

$\boxed{\text{The expected value of } U \text{ is } 200}$

For Z, we have to calculate expected value of Z using Mgf.

$Z = \frac{1}{2}N(100, \theta^2=100) + \frac{1}{2}N(300, \theta^2=100)$

Let $A = N(100, \theta^2=100)$

$B = N(300, \theta^2=100)$

$Z = \frac{1}{2}A + \frac{1}{2}B$

$\phi_Z(t) = \frac{1}{2}E[e^{tA}] + \frac{1}{2}E[e^{tB}]$

$= \frac{1}{2}\phi_A(t) + \frac{1}{2}\phi_B(t)$

Now, we have to find each $\phi_A(t)$, $\phi_B(t)$ and A and B are normal distributions.

$$\Phi_A(t) = \int_{-\infty}^{\infty} e^{tA} \cdot \frac{1}{\sqrt{2\pi\theta_A^2}} e^{-1/2\,(A-\mu_A)^2/\theta_A^2}\, dA$$

$$= \int_{-\infty}^{\infty} e^{tA} \cdot \frac{1}{\sqrt{2\pi\theta_A^2}} e^{-1/2\left(\frac{A-\mu_A}{\theta_A}\right)^2} dA$$

set variable $g = \frac{A-\mu_A}{\theta_A} \rightarrow A = \theta_A g + \mu_A$

$$= \int_{-\infty}^{\infty} e^{t(\theta_A g + \mu_A)} \cdot \frac{1}{\sqrt{2\pi\theta_A^2}} e^{-\frac{1}{2}g^2} dg$$

$$= e^{t\mu_A} \int_{-\infty}^{\infty} e^{t\theta_A g} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}g^2} dg$$

$$= e^{t\mu_A} \cdot e^{\frac{1}{2}\theta_A^2 t^2}$$

Similarly,

$$\Phi_B(t) = e^{t\mu_B} \cdot e^{\frac{1}{2}\theta_B^2 t^2}$$

Therefore,

$$\Phi_Z(t) = \frac{1}{2}\left(e^{t\mu_A} e^{\frac{1}{2}\theta_A^2 t^2}\right) + \frac{1}{2}\left(e^{t\mu_B} e^{\frac{1}{2}\theta_B^2 t^2}\right)$$

If we plug-in $\mu_A = 100$, $\theta_A^2 = 100$, $\mu_B = 300$, $\theta_B^2 = 100$, we get

$$\Phi_Z(t) = \frac{1}{2}\left(e^{100t + 50t^2}\right) + \frac{1}{2} e^{300t + 50t^2}$$

To find the expected value, we have to do a derivative of $\Phi_Z(t)$ with respect to t and set $t=0$

$$\frac{d}{dt}\left(\frac{1}{2}e^{100t+50t^2} + \frac{1}{2}e^{300t+50t^2}\right)$$

$$= \frac{1}{2}(100 + 100t)\, e^{100t+50t^2} + \frac{1}{2}(300 + 100t)\, e^{300t + 50t^2} \Big|_{t=0}$$

$$= \frac{1}{2}(100)\, e^0 + \frac{1}{2}(300)\, e^0$$

$$= 200$$

$\boxed{\text{The expected value of } Z \text{ is } 200}$

③ 3.

$U = \frac{1}{2}(X+Y)$

$Var(U) = Var(\frac{1}{2}(X+Y))$

$= (\frac{1}{2})^2 \; Var(X+Y)$

$= 1/4\,(Var(X) + Var(Y) + Cov(X,Y))$

$Var(X) = \sigma_x^2 = 10^2 = 100$

$Var(Y) = \sigma_Y^2 = 10^2 = 100$

$Cov(X,Y) = E(XY) - E(X)E(Y)$

Since X and Y are independent, $E(XY) = E(X)E(Y)$

$= E(X)E(Y) - E(X)E(Y)$

$= 0$

$Var(U) = 1/4\,(100 + 100 + 0)$

$= 50$

$\boxed{\text{The variance of } U \text{ is } 50}$

To calculate variance of Z, we can calculate it as

$Var(Z) = E(Z^2) - E(Z)^2$

From part 2, we found that $E(Z) = 200$.

Now, we have to find $E(Z^2)$, which is the second derivative of $\phi_Z(t)$, the Mgf function, and plugging 0 for t.

From part 2, we found that the first derivative of $\phi_Z(t)$ is

$\frac{1}{2}(100+100t)e^{100t+50t^2} + \frac{1}{2}(300+100t)\,e^{300t+50t^2}$

$= (50+50t)\,e^{100t+50t^2} + (150+50t)\,e^{300t+50t^2}$

If we take the derivative of the folloLing, we get

$$50e^{100t+50t^2}+(50+50t)(100+100t)e^{100t+50t^2}+50e^{300t+50t^2}+(150+50t)e^{300t+50t^2}(300+100t)$$

If we plug-in $t=0$, we get

$$50+(50)(100)+50+(150)(300)=50+5000+50+45000$$
$$=50100$$

Therefore, $E(Z^2)=50100$.

$$Var(Z)=E(Z^2)-E(Z)^2$$
$$=50100-200^2$$
$$=50100-40000$$
$$=10100$$

$\boxed{\text{The variance of } Z \text{ is } 10100.}$