→ Shreyas Jagdeep Shete (1001888859).

1) Molecular Biology/Genetics Question.

1.1) Mendelian Genetics:—

a)

Ans:—

	RY	Ry	rY	ry
ry	RrYy	Rryy	rrYy	rryy
ry	RrYy	Rryy	rrYy	rryy
ry	RrYy	Rryy	rrYy	rryy
ry	RrYy	Rryy	rrYy	rryy

b) Round Yellow = 4/16

c) Round Green = 4/16

d) Wrinkled Yellow = 4/16

e) Wrinkled Green = 4/16

1.2) Genome Wide Association Studies (GWAS):-

a)

Ans:-

We know that,

$$\text{Bonferroni Correction} = \frac{\text{Significance value}}{\text{number of tests}} = \frac{\alpha}{n}$$

here, $\alpha = 0.05$ and $n = 2.6$ million.

$$\therefore \text{Bonferroni correction} = \frac{0.05}{2.6 \text{ million}} = 1.9230769 2e-8$$

b)

Ans:- Here $\alpha = 0.05$ and $n = 1000$

$$\therefore \text{Bonferroni correction} = \frac{\alpha}{n} = \frac{0.05}{1000} = 0.00005$$

c)

Ans:- Bonferroni Correction is noted to be conservative. That means that, although it protects from Type I error, it is vulnerable to Type II error (i.e., failing to reject the Null hypothesis when

you should in fact reject the null hypothesis

2) Statistics Questions:-

2.1) Drug Approval:-

a) Null hypothesis:- The drug lowers the blood pressure of patients.

$$H_0 : \mu_0 = \mu = 115$$

Alternative hypothesis:-
The drug doesnot lower the blood pressure of patients.

$$H_a : \mu \neq \mu_0 \text{ i.e., } \mu \neq 115$$

b) Here since the sample size $n = 20$ which is less than 30 and standard deviation is given, it is going to be Z-test that is appropriate to test statistical significance.

We know that

$$Z = \frac{(\bar{X} - \mu_0)}{S} = \frac{(\bar{X} - \mu_0)}{\sigma/\sqrt{n}}$$

Here, it's given that $\mu_0 = 120$.

$\bar{X} = 115$

$\sigma = 15$ and $n = 20$. degree of freedom $= n - 1 = 19$

$$\Rightarrow Z_{19} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{115 - 120}{15/\sqrt{20}} = \frac{-5}{15/\sqrt{20}}$$

$$\therefore Z_{19} = -1.490711985 \Rightarrow |Z_{19}| = 1.490711985$$

Given $\alpha = 0.05$

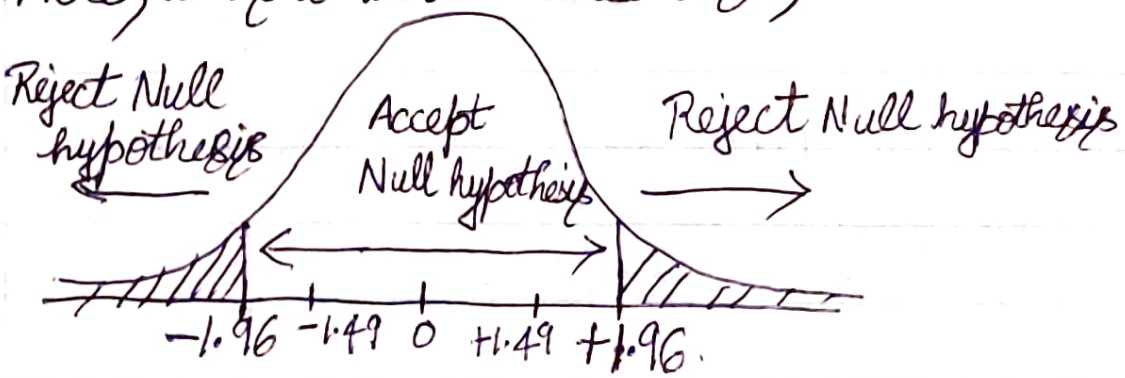
$\therefore Z_{critical} = Z_{\alpha/2} = Z_{0.025} = 1.96$.

$p\text{-value} = p[|Z| < 1.96] \Rightarrow$ pvalue $= 0.06$.

$\Rightarrow 1.4907 < 1.96$. Also, pvalue $> \alpha \Rightarrow 0.06 \geq 0.05$

(c) $\therefore$ Null hypothesis H_0 is valid and accepted.

Here, it is a two-tailed test,



∴ The drug lowers the blood pressure of patients.

Yes, this drug can be put on the market.

The sample of 20 mice taken for experiment will have similar result in the population and that in human beings is a factor considered in making the decision.

2.2)

a) The advantage of a non-parametric test vs a parametric test are:-

1) More statistical power when assuming for the parametric tests have been violated. When assumptions haven't been violated, they can be almost as powerful.
2) Fewer assumptions (i.e., the assumption of normally doesn't apply).
3) Small sample sizes are accepted.
4) They can be used for all data types

b) As non-parametric test makes fewer assumptions, it has less powerful than parametric test in distinguishing between

the two arms or variance in distribution.

4) Difficulty Adjustment

a) It took me more than a day to complete it.

b) For me, as I have windows and not MAC I could not or took longer time to set up and code for the assignment. As, I am new to python I was expecting a little basics to be covered before moving to harder problem.