No. ______

Date ______

BELER, DOMINIC EMIL R. CAS-05-601A

MILITANTE, DARIEL M. BAYESIAN STATISTICS

1.) (a) Upon opening the nips bag, a categorical data can be identified since the context is whether ~~a~~ it is blue or not. Though, it can be quantified by assigning exactly two numerical values for ~~the~~ each of the categorical data (blue or not blue).

(b) One way to estimate the percentage of blue is to determine its number given that all colors, which is 5 in total, have ~~even~~ equal chances to be in a nips bag. However, the pitfall is that the distribution of colors ~~is~~ might be random in the factory. This is on top of the error in which the no. of nips in a bag is not exactly the same all the time.

(c) Given that there is no blue nips in a bag, assuming that the entire population also has zero blue nips is not realistic at all since again, the distribution might be random.

(d) By actually counting the number of blue and non-blue nips in a bag; non-blue meaning all other colors ~~be~~ except blue.

2.) ~~Posterior~~ = $\frac{(\text{Likelihood})(\text{Prior})}{\text{marginal}}$ or ~~mathematically:~~

(a) $P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$; A = getting a blue nip, B = getting a non-blue nip

3) (a) The likelihood can therefore be calculated via Binomial distribution formula given as:

$$f(x) = nCr\,(p)^{r}\,(1-p)^{n-r};$$

n = total no. of nips in a bag

r = no. of blue nips in a bag

p = probability of getting blue in a bag

f(x) = likelihood of getting non-blue given you've already got blue

(b) w/o replacement since ~~Baye's rule talks about~~ independence, given we didn't return any nips ~~in the bag~~.

(c) Yes, we can eat it already if the data is recorded. ~~data~~

No. ______
Date ______

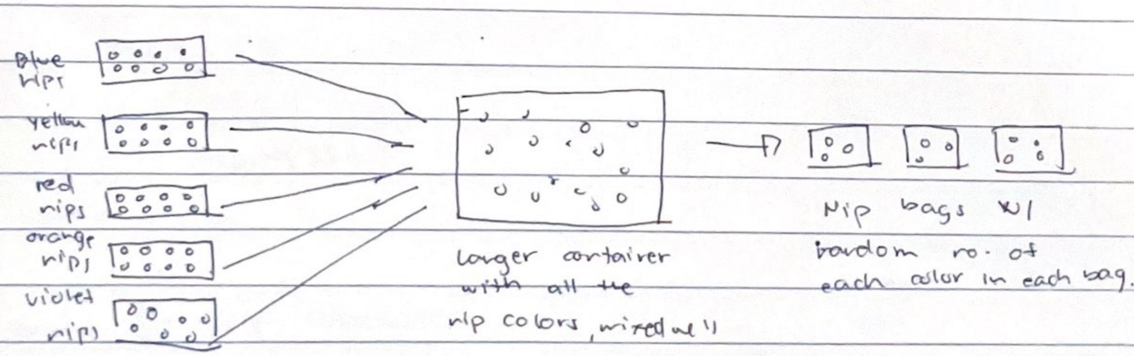
4.) ⓐ The prior information can be approximated by getting the probability of blue nips observed upon opening a random bag of nips.

ⓑ We assume that each color of nips are made separately ~~then~~ before putting it all in a larger container. ~~After~~ Only after the the nips are packed in a small bag.

ⓒ Yes. The manufacturing should affect the proportions of blue nips in a bag given that the no. of blue nips is not fixed in a bag. (This is on top of the error that the total nips in a bag is not also fixed)

ⓓ. Yes. The nips ~~a~~ should be well-mixed befor packaging in a bag since the proportions of blue nips is not evenly distributed in all the bags, meaning the packaging should be random.

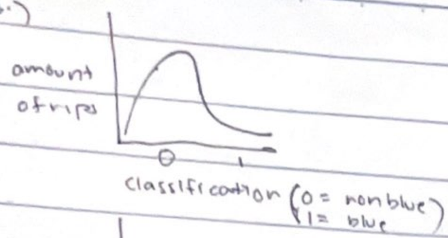
ⓔ Realistically, we can never identify the population of blue nips in a factory. However, we have a theory that all colors are made equal in number, but before packaging in a bag, it will be put in a larger container first and mixed well.

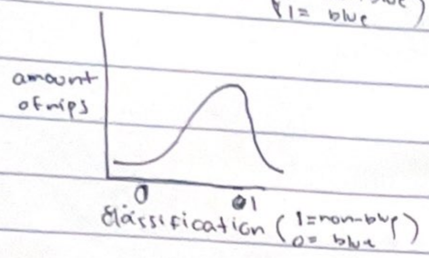


No. ________
Date ________

5.)

(c)





(a)

Most of the prior probability should fall to where the blue nips is distributed since this is our probability of reference.

(c)

Since in our bag, there is fewer blue than non-blue, thus the low probability should also be ~~in the~~ where the blue distribution resides.

6) FINAL QUESTION:

Can we actually predict the no. of blue nips in the 2nd bag, given the blue nips in the 1st bag?

$$\text{posterior} = \frac{(\text{likelihood})(\text{prior})}{\text{marginal}}$$

Given:

	1st pack	2nd pack
Blue	2	1
NonBlue	10	10
Total	12	11

likelihood:

$$f(x) = {}_nC_r \, (p)^r (1-p)^{n-r}$$

$n = 12$, $r = 2$, $p = \frac{2}{12}$ or $\frac{1}{6}$

$$= {}_{12}C_2 \left(\tfrac{1}{6}\right)^2 \left(1-\tfrac{1}{6}\right)^{(12-2)}$$

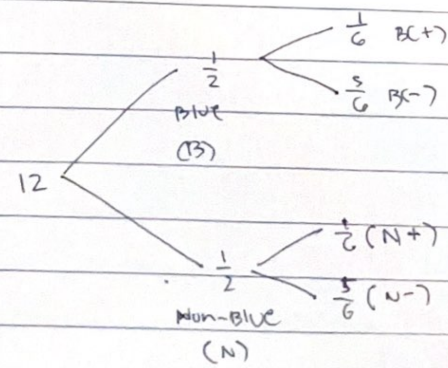
$$= {}_{12}C_2 \left(\tfrac{1}{6}\right)^2 \left(\tfrac{5}{6}\right)^{10}$$

$$= 0.2960935685$$

$f(x) = 0.296$

prior = $\frac{2}{12}$ or $\frac{1}{6}$ (observed)

marginal:



marginal = $P(B+|B) \cdot P(B) + P(N-|N)\,P(N)$

$= \left(\frac{1}{6} \cdot \frac{1}{2}\right) + \left(\frac{5}{6} \cdot \frac{1}{2}\right)$

$= \frac{1}{12} + \frac{5}{12}$

marginal $= \frac{1}{2}$

* continuation at back

No. ______

Date ______

$$\cancel{P(A|B)} = \frac{\overset{\text{likelihood}}{\cancel{(\text{posterior})}}\,(\text{prior})}{\text{morginal}}$$

$$= \frac{(0.296)\left(\frac{1}{6}\right)}{\frac{1}{2}}$$

$$= 0.098667$$

$$= 0.099$$

$$P(A|B) = 9.9\%$$

2ND PACK:

observed: 1 blue, 10 non-blue

$$P(\text{Blue}) = \frac{1}{11}$$

$$= 0.091$$

$$= 9.1\%$$

Although the posterior probability (9.9%) is close to the probability of getting blue in 2nd pack (9.1%), this would not hold true to other packs with more blue nips or without any blue nips at all. Again, this is because we have an assumption that the proportion of blue in a pack is randomly distributed because of the nature of the manufacturing in the factory.

No. ______

Date ______

$$P(A|B) = \frac{(\text{~~posterior~~ likelihood})(\text{prior})}{\text{marginal}}$$

$$= \frac{(0.296)\left(\frac{1}{6}\right)}{\frac{1}{2}}$$

$$= 0.098667$$

$$= 0.099$$

$$P(A|B) = 9.9\%$$

2ND PACK:

observed: 1 blue, 10 non-blue

$$P(\text{Blue}) = \frac{1}{11}$$

$$= 0.091$$

$$= 9.1\%$$

Although the posterior probability (9.9%) is close to the probability (9.1%) of getting blue in 2nd pack, this would not hold true to other packs with more blue nips or without any blue nips at all. Again, this is because we have an assumption that the proportion of blue in a pack is randomly distributed because of the nature of the manufacturing in the factory.