STAT 545 Homework 1

Yinsen Miao [*]

September 7, 2015

Problem 1

(0.5 point each)

(a) UK political party preference: nominal

(b) Anxiety rating: ordinal

(c) Patient survival: interval

(d) Clinical location: nominal

(e) Response of tumor to chemotherapy: ordinal

(f) Favorite beverage: nominal

(g) Appraisal of company's inventory level: ordinal

Problem 2

(a) (1 point) Y has a binomial distribution, since:

$$\begin{aligned} Var(Y) &= Var\left(\sum_i Y_i\right) \\ &= \sum_i Var(Y_i) \\ &= n\pi(1-\pi) \end{aligned}$$

where $Y_i \sim Bernouli(\pi)$.

(b) (2 points) Since Y_is are not independent, (a) doesn't hold any more

[*]Department of Statistics, Rice University, E-mail ym11@rice.edu

$$
\begin{aligned}
Var(Y) &= Var\left(\sum_i Y_i\right) \\
&= E\left[\sum_i Y_i\right]^2 - \left[E\left(\sum_i Y_i\right)\right]^2 \\
&= E\left[\sum_i \sum_j Y_i Y_j\right] - \sum_i \sum_j EY_i EY_j \\
&= \sum_i \sum_j \left(EY_iY_j - EY_iEY_j\right) = \sum_i \sum_j Cov\left(Y_i, Y_j\right) \\
&= \sum_i Var\left(Y_i\right) + \sum_{i \neq j} Cov\left(Y_i, Y_j\right) \\
&> n\pi\left(1-\pi\right)
\end{aligned}
$$

(c) (2 points) By the Eve's Law,

$$
\begin{aligned}
Var(Y) &= Var\left(E\left[Y \mid \pi\right]\right) + E\left[Var\left(Y \mid \pi\right)\right] \\
&= Var\left(n\pi\right) + E\left[n\pi\left(1-\pi\right)\right] \\
&= n^2 Var\left(\pi\right) + nE\left[\pi\right] - n\left(E\left[\pi^2\right] + Var\left(\pi\right)\right) \\
&= nE\left[\pi\right]\left(1 - E\left[\pi\right]\right) - n\left(n-1\right) Var\left[\pi\right] \\
&= n\rho\left(1-\rho\right) + n\left(n-1\right) Var\left[\pi\right] \\
&> n\rho\left(1-\rho\right)
\end{aligned}
$$

(d) (2 points)

$$
\begin{aligned}
P\left(Y_i = 1, Y_j = 1\right) &= \int\int P\left(Y_i = 1, Y_j = 1 \mid \pi_i, \pi_j\right) g\left(\pi_i\right) g\left(\pi_j\right) d\pi_i d\pi_j \\
&= \int\int \pi_i \pi_j g\left(\pi_i\right) g\left(\pi_j\right) d\pi_i d\pi_j \\
&= \int \pi_i g\left(\pi_i\right) d\pi_i \int \pi_j g\left(\pi_j\right) d\pi_j \\
&= E\left[\pi_i\right] E\left[\pi_j\right] = \rho^2 = P\left(Y_i = 1\right) P\left(Y_j = 1\right)
\end{aligned}
$$

which indicates $\{Y_i\}$s are marginally independent and $P\left(Y_i = 1\right) = \rho$. Hence, $Y = \sum_i Y_i$ follows $Binomial\left(n, \rho\right)$. However, when conditionally on $\{\pi_i\}$, Y becomes a sum of independent Bernoulli trials with unequal probabilities (You also can formally show this by calculating its characteristic function and show it is not one of a binomial distribution).