HW1 STT 465
(MSU, Fall, 2015)

Due Wednesday, Sept 14th in class.

Topics: marginal and conditional probability, Bayes theorem, exchangeability.

Background. Consider a disease controlled by a single DNA locus with two possible alleles A and B. In this setting any individual in the population can have one of the three following genotypes: AA, AB and BB.

From a large genetic epidemiology study the following are the estimated probability of disease $p(D_i=1 \mid G_i)$ for each of the possible genotypes

Genotype (G_i)	$p(D_i=1 \mid G_i)$
AA	0.0
AB	0.1
BB	0.9

If θ denotes the frequency of allele A, under random mating, the genotype frequencies in the population are: $p(AA)= \theta^2$, $p(BB)= (1-\theta)^2$, $p(AB)= 2\theta(1-\theta)$

Question 1

In population I $\theta=0.9$, compute and report the prevalence of disease in population I.

Question 2

An individual in population I has developed the disease, what is the probability that the genotype of that individual is AA, AB or BB?

Question 3

In population II the frequency of allele A is 0.95. An individual is healthy, what is the probability that this individual comes from population I (assume that, a priori, individuals are equally likely to come from population 1 or 2).

Question 4.

Provide frequencies for the joint distribution of the two Bernoulli random variables (i.e., probabilities in a 2x2 contingency table) that:

- satisfy IID (identically and independently distributed),
- satisfy exchangeability but not IID
- do not satisfy IID and are not exchangeable.

Explain your reasoning.