Answer Key Spring 2018, MA/STAT 519 Midterm Exam

1. Consider the matching-hat-problem for n people with n hats. All the people throw their hats into a container and then each of them picks one at random. Let $P_n(0)$ be the probability that no people get their own hats back. In class, it was derived that

$$P_n(0) = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}.$$

(Ross. p.93 eg 5d)

Now let $P_n(k)$ be the probability of exactly k ($k \leq n$) people get their own hats back. Prove that

$$P_n(k) = \frac{1}{k!}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{(n-k)}}{(n-k)!}\right).$$

(Hint: consider two groups of people separately, one with all people with matching hats and the other with no people with matching hats.)

n people, n hats

choose k people

(n-k) remaining people

$$\binom{n}{k} \underbrace{\frac{1}{n}\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)\cdots\left(\frac{1}{n-k+1}\right)} P_{n-k}(0).$$

Note: The k people choose among n hats!

$$\frac{\cancel{n!}}{k!\,\cancel{(n-k)!}} \cdot \frac{1}{\cancel{n(n-1)(n-2)\cdots(n-k+1)}} P_{n-k}(0)$$

$$= \frac{1}{k!}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{n-k}}{(n-k)!}\right)$$

(Ross (Self Test Problems) p. 173 #4.)

2. A certain community consists of m families. For $i = 1, 2, \ldots r$, let n_i be the number of families with i children so that $\sum_{i=1}^{r} n_i = m$ and $\sum_{i=1}^{r} i n_i$ gives the total number of children in the community.

   If one of the families is randomly chosen, let X be the number of children in that family. Alternatively, if one of the children in the whole community is randomly chosen, let Y be the total number of children in the family of that child.

   Which of $E(X)$ and $E(Y)$ is bigger? Prove your statement rigorously, i.e. mathematically.

   (Hint: simplify, expand, and compare terms.)

$$EX = \sum_{i=1}^{r} i P(X=i) = \sum_{i=1}^{r} i \frac{n_i}{m}$$

$$= \left(\frac{\sum_{i=1}^{r} i n_i}{m} \right)$$

$$EY = \sum_{i=1}^{r} i P(Y=i) = \sum_{i=1}^{r} i \left(\frac{i n_i}{\sum_{j=1}^{r} i n_j} \right)$$

$$= \frac{\sum_{i=1}^{r} i^2 n_i}{\left(\sum_{j=1}^{r} i n_j \right)}$$

$$\boxed{EY > EX}$$

Because children from larger families are more likely to be picked

$$EY > EX \quad ?$$

You can use this blank page.

$$\frac{\sum_i i^2 n_i}{\sum_i i n_i} \overset{?}{>} \frac{\sum_i i n_i}{\sum_i n_i}$$

(m)

$$\iff \left(\sum_i i^2 n_i\right)\left(\sum_j n_j\right) \overset{?}{>} \left(\sum_i i n_i\right)^2$$

$$\iff \cancel{\sum_{i=j} i^2 n_i^2} + \sum_{i\neq j} i^2 n_i n_j \overset{?}{>} \cancel{\sum_{i=j} i^2 n_i^2} + \sum_{i\neq j} ij\, n_i n_j$$

$$\iff \sum_{i<j} (i^2+j^2) n_i n_j \overset{?}{>} \sum_{i<j} 2ij\, n_i n_j$$

$$\iff i^2+j^2 \overset{?}{>} 2ij$$

Yes, since $(i-j)^2 > 0$

$$\left(i^2+j^2 > 2ij \right).$$

(Ross. p.101, # 46
Homework #3)

3. In any given year, a male automobile policyholder will make a claim with probability p_M and a female automobile policyholder will make a claim with probability p_F. Let α be the fraction of male policyholders (and $(1-\alpha)$ be the fraction of female policyholders). Now a policyholder is randomly chosen and fixed for the following consideration. Let A_i denote the event that this person makes a claim during the i-th year.

   (a) Suppose $p_M \neq p_F$. Prove that $P(A_2|A_1) > P(A_1)$. Are A_1 and A_2 independent?

   (b) Suppose $p_M > p_F$. Let M be the event that the policyholder is a male. Prove that $P(M|A_1) > P(M)$.

(a) $$P(A_2|A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$$

$$= \frac{P(A_2 \cap A_1 | M)P(M) + P(A_2 \cap A_1 | F)P(F)}{P(A_1|M)P(M) + P(A_1|F)P(F)}$$

$$= \frac{p_M^2 \alpha + p_F^2(1-\alpha)}{p_M \alpha + p_F(1-\alpha)} \overset{?}{>} p_M\alpha + p_F(1-\alpha)$$

$$\Longleftrightarrow p_M^2\alpha + p_F^2(1-\alpha) \overset{?}{>} \left(p_M\alpha + p_F(1-\alpha)\right)^2$$

$$\Longleftrightarrow p_M^2(\alpha - \alpha^2) + p_F^2((1-\alpha) - (1-\alpha)^2) - 2p_Mp_F\alpha(1-\alpha) > 0$$

$$\Longleftrightarrow p_M^2(\alpha\cancel{(1-\alpha)}) + p_F^2\cancel{(1-\alpha)}(\alpha) - 2p_Mp_F\alpha\cancel{(1-\alpha)} > 0$$

$$\Longleftrightarrow (p_M - p_F)^2 > 0 \quad \checkmark$$

A_1, A_2 are not independent as $P(A_2|A_1) \neq P(A_1)$

You can use this blank page.

(b) $P(M \mid A_1) \overset{?}{>} P(M)$

$\iff \dfrac{P(M \cap A_1)}{P(A_1)} > P(M)$

$\iff \dfrac{P(A_1 \mid M)\cancel{P(M)}}{P(A_1)} > \cancel{P(M)}$

$\iff p_M > P(A_1)$

$\iff p_M > p_M\alpha + p_F(1-\alpha)$

$\iff p_M\cancel{(1-\alpha)} > p_F\cancel{(1-\alpha)}$

$\iff p_M > p_F$ ✓

4. Let X be a positive-integer-valued random variables, i.e. it takes values in $\{1, 2, 3, \ldots\}$. X is said to enjoy the memoryless property if for any positive integers m, n, it holds that

$$P(X > m+n | X > m) = P(X > n).$$

Prove that X enjoys the memoryless property *if and only if* X is a geometric random variable, i.e. there is a p (with $0 < p < 1$) such that $P(X = i) = p(1-p)^{i-1}$.

Note: you need to prove *two* separate statements: geometric implies memoryless *and* memoryless implies geometric. For simplicity, just ignore the "degenerate" case $p = 0$ or $p = 1$.

$$P(X > m+n / X > m) = \frac{P(X > m+n, X > m)}{P(X > m)}$$

$$= \frac{P(X > m+n)}{P(X > m)}$$

Hence memoryless property is equivalent to:

$$P(X > m+n) = P(X > n) P(X > m)$$

① Geometric $\Longrightarrow$ memoryless.

$$P(X > n) = P(\text{1st } n \text{ trials give } 0)$$

$$= q^n$$

$$\Big(\text{for, } P(X > n) = \sum_{i=n+1}^{\infty} P(X = i) = \sum_{i=n+1}^{\infty} p q^{i-1}$$

$$= p q^n \sum_{j=0}^{\infty} q^j = \cancel{p} q^n \frac{1}{\cancel{1-q}} = q^n\Big)$$

Hence, similarly, $P(X>m)=q^m$

You can use this blank page.

$P(X>m+n)=q^{m+n}$

~~Hence~~ note: $q^{m+n}=q^n\cdot q^m$

Hence memoryless property holds.

(2) Memoryless $\Rightarrow$ geometric

$P(X>m+n)>P(X>m)P(X>n)$ $\quad n=1+(n-1)$

In particular, $P(X>n)=P(X>1)P(X>n-1)$

$=P(X>1)^2P(X>n-2)$

$=P(X>1)^3P(X>n-3)$

$\vdots$

$=P(X>1)^{n-1}P(X>1)$

$=P(X>1)^n$

$$\boxed{\begin{aligned}&\text{Let } p=P(X=1)\\&\text{Then } P(X>1)=1-P(X=1)\\&\qquad\qquad\qquad =1-p\end{aligned}}$$

$P(X=n)=P(X>n-1)-P(X>n)$

$=(1-p)^{n-1}-(1-p)^n=(1-p)^{n-1}(1-1+p)$

$=(1-p)^{n-1}p$

5. There are two types ($i = 1, 2$) of coins in a bin, each with probability p_i to give a head upon tossing. Let α ($0 < \alpha < 1$) be the fraction of type 1 coins. A coin is randomly chosen and is repeatedly tossed. Suppose the first m tosses give all tails. What is the probability that the next n tosses still give all tails.

Let X be the # of trials corresponding to the first head. X is a geometric r.v.

(* Given a known value of p)

$$P(X > m+n \mid X > m)$$

(* Note cannot use memoryless property here as p is not known.)

$$= \frac{P(X > m+n,\ X > m)}{P(X > m)}$$

$$= \frac{P(X > m+n)}{P(X > m)}$$

$$= \frac{P(X > m+n \mid \text{coin 1})\,P(\text{coin 1}) + P(X > m+n \mid \text{coin 2})\,P(\text{coin 2})}{P(X > m \mid \text{coin 1})\,P(\text{coin 1}) + P(X > m \mid \text{coin 2})\,P(\text{coin 2})}$$

$$= \frac{(1-p_1)^{m+n}\alpha + (1-p_2)^{m+n}(1-\alpha)}{(1-p_1)^{m}\alpha + (1-p_2)^{m}(1-\alpha)}$$