Ph. D. Qualifying Exam: Applied Probability. June 1st, 2017

Instructor: Bei Hu **Candidate:**____

There are 8 problems. Show all your work.

1. Toss a fair coin repeatedly until two consecutive heads appear. Let X be the total of all head count, find EX, $VarX$. Here are some sample events (1 = head, 0 = tail):

$$\{1,1\}, X = 1+1 = 2;$$
$$\{1,0,1,1\}, X = 1+0+1+1 = 3;$$
$$\{1,0,0,0,1,1\}, X = 1+0+0+0+1+1 = 3;$$
$$\{1,0,0,1,0,1,1\}, X = 1+0+0+1+0+1+1 = 4;$$

Sol. Last two toss would have to be heads if $X = 2$, thus

$$P(X=2) = \sum_{j=2}^{\infty} 2^{-j} = 1/2.$$

For $k > 2$,

$$P(X=k) = P(\text{1st toss is tail})P(X=k) + P(\text{1st toss is head})P(\text{2nd toss is tail})P(X=k-1),$$

so that

$$\frac{1}{2}P(X=k) = \frac{1}{4}P(X=k-1).$$

By induction $P(X=k) = 2^{-k+1}$. Thus

$$EX = \sum_{k=2}^{\infty} k 2^{-k+1} = 3,$$
$$EX^2 = \sum_{k=2}^{\infty} k^2 2^{-k+1} = 11,$$
$$VarX = EX^2 - (EX)^2 = 2.$$

2. Suppose X and Y are independent continuous random variables with the same density function $f(x) = e^{-x},\ x > 0$.

(a) Find the density function for $X + Y$;

Sol.

$$\int_0^x e^{-u} e^{-(x-u)} du = xe^{-x}, \qquad x > 0.$$

(b) Find the joint density function for $X, X + Y$.

Sol.

$$\begin{aligned} P(X \le u, X + Y \le v) &= \int\int_{x \le u,\ x+y \le v} e^{-x-y} dx dy \\ &= \int_0^u \Big(\int_0^{v-x} e^{-x-y} dy \Big) dx \\ &= \int_0^u (e^{-x} - e^{-v}) dx \end{aligned}$$

Taking derivatives in u then in v, we obtain

$$f_{X,X+Y}(u, v) = e^{-v}, \qquad 0 < u < v.$$

3. Three factories A, B and C manufacture zoogles. Factory A produces 20% of the zoogles and factory B produces 75% of the zoogles. The remaining 5% of the zoogles are from factory C. The defective rate for factory A is 1 in 50 and the defective rate for factory B is 1 in 20. The defective rate for factory C is 1 in 100.

(a) What is the probability a randomly selected zoogle is defective?

Sol.

$$\begin{aligned} P(\text{def}) &= P(\text{from A})P(\text{def from A}) + P(\text{from B})P(\text{def from B})P(\text{from C})P(\text{def from C}) \\ &= 0.2 \cdot \frac{1}{50} + 0.75 \cdot \frac{1}{20} + 0.05 \cdot \frac{1}{100} \\ &= 0.042 \end{aligned}$$

(b) Given that a zoogle is defective, what is the probability that it is from factory C?

Sol.

$$\begin{aligned} P(\text{from C}|\text{def}) &= P(\text{from C and def})/P(\text{def}) \\ &= \frac{0.05 \cdot \frac{1}{100}}{0.042} = 0.0119 \end{aligned}$$

(c) A shipment of 100 zoogles is received from a distribution center (which distribute zoogles from all factories). What is the probability that all of them are not defective?

Sol.

$$Prob = (1 - 0.042)^{100} = 0.958^{100} = 0.01369$$

4. Here is a simple random walk. Let $S_n = S_0 + \sum_{i=1}^{n} X_i$, where X_i are independent, and takes the value $-1, 1$. Assume

$$P(X_i = 1) = p, \ P(X_i = -1) = q, \quad p + q = 1, \quad p > q.$$

Suppose that the sites 0 and N are absorbing barriers. Let T_k be the time it takes to go from the site k to one of the absorbing barriers. Let $J_k = E(T_k)$. Obviously $J_0 = 0$ and $J_N = 0$.

(a) Find the equation for J_k.

Sol. For $1 \le k \le N-1$,

$$\begin{aligned} J_k &= P(\text{1st step is right})(J_{k+1}+1) + P(\text{1st step is left})(J_{k-1}+1) \\ &= pJ_{k+1} + qJ_{k-1} + 1, \end{aligned}$$

or

$$J_k - pJ_{k+1} - qJ_{k-1} = 1.$$

(b) Find J_k.

Sol. Inhomogeneous particular solution $J_k = Ak$:

$$Ak - pA(k+1) - qA(k-1) = 1, \qquad A(q-p) = 1, \qquad A = -\frac{1}{p-q}, \qquad J_k = -\frac{k}{p-q}.$$

Homogeneous solution $J_k = \theta^k$:

$$\theta^k - p\theta^{k+1} - q\theta^{k-1} = 0, \qquad \theta - p\theta^2 - q = 0, \qquad \theta = 1, q/p$$

General solution

$$J_k = C_1 + C_2(q/p)^k - \frac{k}{p-q}.$$

The boundary condition $J_0 = 0$ implies $C_1 + C_2 = 0$, the boundary condition $J_N = 0$ implies $C_1 + C_2(q/p)^N - N/(p-q) = 0$. Thus

$$-C_2 = C_1 = \frac{N}{(p-q)(1-(q/p)^N)}$$

and

$$J_k = \frac{N\big(1-(q/p)^k\big)}{(p-q)(1-(q/p)^N)} - \frac{k}{p-q}.$$

5. Suppose $G(s)$ is the generating function of the random variable X. Suppose also that (a) the random variable X takes positive integer values, and (b) X_j ($j = 1, 2, 3, \cdots$) have the same distribution as X, and (c) $X, X_1, X_2, \cdots$ are all independent.

(a) Find $E\left(\frac{1}{X}\right)$ in terms of the function G (and possibly derivatives or integrals of G).

$$E\left(\frac{1}{X}\right) = \int_0^1 s^{X-1} ds = \int_0^1 \frac{G(s)}{s} ds.$$

(b) Let $Y = X_1 + X_2 + X_3 + \cdots X_X$. Find $E(Y)$ in terms of the function G (and possibly derivatives or integrals of G).

Sol. The generating function of Y is given by $G(G(s))$, so that

$$EY = \frac{d}{ds} G(G(s))\Big|_{s=1} = G'(G(1))G'(1) = (G'(1))^2.$$

6. (a) Suppose that $S_n = X_1 + X_2 + \cdots X_n$ is a Martingale with respect to $X_1, X_2, \cdots, X_n$, i.e.,

$$E(S_{n+1}|X_1, \cdots, X_n) = S_n.$$

Prove that $E(X_iX_j) = 0$ if $i \neq j$.

Sol. Suppose $j > i$. Then

$$\begin{aligned} E(X_iX_j) &= E(X_i(S_j - S_{j-1}) \\ &= E\Big(E(X_i(S_j - S_{j-1})|X_1, \cdots, X_{j-1})\Big) \\ &= E\Big(X_iE((S_j - S_{j-1})|X_1, \cdots, X_{j-1})\Big) \\ &= 0. \end{aligned}$$

(b). For $x > 0$, find the limit (quote explicitly any theorem you use)

$$\lim_{n\to\infty} 2^{-n} \sum_{k:|k-\frac{1}{2}n|\leq\frac{1}{2}x\sqrt{n}} \binom{n}{m}$$

Sol. For Bernoulli distribution with $p = q = 1/2$, $EX_i = p = 1/2$, $VarX_i = p(1-p) = 1/4$. Let $S_n = \sum_{i=1}^n X_i$. By central limit theorem,

$$\begin{aligned} \lim_{n\to\infty} 2^{-n} \sum_{k:|k-\frac{1}{2}n|\leq\frac{1}{2}x\sqrt{n}} \binom{n}{m} &= \lim_{n\to\infty} \sum_{k:|k-ES_n|/\sqrt{VarS_n}\leq x} P(S_n = k) \\ &= \int_{-x}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \end{aligned}$$

7. Let W_t be the standard Brownian motion with $W_0 = 0$, $EW_t = 0$ and $VarW_t = t$. Suppose that $0 < s < t$. Find

(a) $E(W_t|W_s)$ and $Var(W_t|W_s)$.

Sol.

$$\begin{aligned} E(W_t|W_s) &= E(W_t - W_s + W_s\big|W_s) \\ &= W_s + E(W_t - W_s\big|W_s) \\ &= W_s + E(W_{t-s}) \\ &= W_s. \end{aligned}$$

$$\begin{aligned} E(W_t^2|W_s) &= E((W_t - W_s + W_s)^2\big|W_s) \\ &= E\Big((W_t - W_s)^2 + 2W_s(W_t - W_s) + W_s^2\big|W_s\Big) \\ &= E\Big((W_t - W_s)^2\big|W_s\Big) + 2W_sE\Big((W_t - W_s)\big|W_s\Big) + W_s^2 \\ &= EW_{t-s}^2 + 2W_sEW_{t-s} + W_s^2 \\ &= t - s + W_s^2. \end{aligned}$$

$$Var(W_t|W_s) = E(W_t^2|W_s) - E(W_t|W_s)^2 = t - s + W_s^2 - W_s^2 = t - s.$$

(b) Using Itô's formula, find $d(W_t)^2$.

Sol. $f(w) = w^2,\ f'(w) = 2w,\ f''(w) = 2.$

$$d(W_t)^2 = 2W_tdW_t + \frac{1}{2} \cdot 2 \cdot dt = 2W_tdW_t + dt$$

(c) Find $\int_0^t W_tdW_t$

Sol.

$$\int_0^t W_tdW_t = \frac{1}{2}\int_0^t d(W_t)^2 - dt = \frac{1}{2}(W_t)^2 - \frac{t}{2}.$$

8. Let W_t be the standard Brownian motion with $W_0 = 0$, $EW_t = 0$ and $VarW_t = t$.

(a) Prove that $\{W_{t_1}, W_{t_2}, \cdots, W_{t_N}\}$ $(t_1 < t_2 < \cdots < t_N)$ satisfies a multivariate normal distribution.

Proof: Since

$$\begin{aligned}&\{W_{t_1} = x_1,\ W_{t_2} = x_2,\ \cdots, W_{t_N} = x_N\}\\ &\quad = \{W_{t_1} = x_1,\ W_{t_2} - W_{t_1} = x_2 - x_1,\ \cdots, W_{t_N} - W_{t_{N-1}} = x_N - x_{N-1}\}\end{aligned}$$

and the components of the second expression are all independent by definition of Brownian motion, it is easy to see that a multivariate normal distribution is generated.

(b) Suppose that $0 < s < t$. Find $E(W_s|W_t)$.

Sol. Since $EW_s = EW_t = 0$,

$$Cov\Big(W_s - \frac{s}{t}W_t, W_t\Big) = E\Big[(W_s - \frac{s}{t}W_t)W_t\Big] = E(W_sW_t) - \frac{s}{t}EW_t^2 = s - \frac{s}{t}\cdot t = 0,$$

The processes $W_s - \frac{s}{t}W_t$ and W_t are uncorrelated. Because Brownian motion generates multivariate normal distribution, they are actually independent. Thus

$$\begin{aligned}E(W_s|W_t) &= E\Big(W_s - \frac{s}{t}W_t + \frac{s}{t}W_t\Big|W_t\Big)\\ &= E\Big(W_s - \frac{s}{t}W_t\Big|W_t\Big) + \frac{s}{t}W_t\\ &= E\Big(W_s - \frac{s}{t}W_t\Big) + \frac{s}{t}W_t\\ &= \frac{s}{t}W_t.\end{aligned}$$