Chase Joyner

901 Final Exam

December 15, 2017

Problem 1

Johnny's favorite cereal is having a contest. There are N distinct coupons and upon collecting all N coupons, you win a prize. Each day Johnny gets a new box of cereal and hence a new coupon. The probability of getting the ith coupon is $1/N$. Let $T_1 = 1$ and for $n = 2, ..., N$, T_n be the day Johnny gets a coupon different from those obtained in days $T_1, ..., T_{n-1}$. Thus T_N is the day that Johnny gets all N coupons.

(a) Show that T_N is the sum of N independent geometrically distributed random variables.

> **Solution:** Let X_i denote the number of the days it takes to obtain the ith unique coupon after having obtained $(i-1)$ unique coupons. Then X_i is geometrically distributed with probability $p_i = (N-i+1)/N$. Since each day is independent, we have the X_i's are independent and clearly $T_N = X_1 + \cdots + X_N$.

(b) Find the expected value and variance of T_N.

> **Solution:** Utilizing the mean of geometric random variables based on trials,
>
> $$E[T_N] = \sum_{i=1}^{N} \frac{N}{N-i+1} = N h_N$$
>
> where h_N is the sum of the first N terms in the harmonic sequence. Now utilizing the variance of geometric random variables, by independence,
>
> $$V(T_N) = \sum_{i=1}^{N} \frac{1-p_i}{p_i^2} = \sum_{i=1}^{N} \frac{Ni - N}{(N-i+1)^2}.$$

(c) Show that $T_N/(N \log N)$ converges in probability to 1.

> **Solution:** First note that
>
> $$V\big(T_N/(N\log N)\big) = \frac{1}{N^2 \log^2 N} \sum_{i=1}^{N} \frac{Ni-N}{(N-i+1)^2} \le \frac{1}{N^2 \log^2 N} \sum_{i=1}^{N} \frac{N^2}{(N-i+1)^2}$$
> $$\le \frac{1}{N^2 \log^2 N} \sum_{i=1}^{N} \frac{N^2}{i^2} \le \frac{C}{\log^2 N}$$

where $C = \sum_{i=1}^{\infty} i^{-2}$. Then, $V\big(T_N/(N \log N)\big) \to 0$ as $N \to \infty$. Now, consider

$$\int_1^N \frac{1}{x} dx < h_N < 1 + \int_1^N \frac{1}{x} dx$$

which is $\log N < h_N < 1 + \log N$. Thus $h_N / \log N \to 1$, implying $E[T_N/(N \log N)] \to 1$. Therefore, for $\epsilon > 0$, there is an N large enough such that

$$\big| E[T_N/(N \log N)] - 1 \big| < \epsilon/2.$$

Thus, for any $\epsilon > 0$,

$$\begin{aligned} P\big(|T_N/(N \log N) - 1| \geq \epsilon\big) &\leq P\big(|T_N/(N \log N) - h_N/\log N| \geq \epsilon/2\big) \\ &\leq \frac{V\big(T_N/(N \log N)\big)}{(\epsilon/2)^2} \to 0 \end{aligned}$$

proving that $T_N/(N \log N) \to 1$ in probability.

Problem 2

Suppose that $X_1, X_2, ...$ are iid having expectation μ and variance σ^2. Let $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ so that $\sqrt{n}(\bar{X} - \mu)/\sigma$ converges in distribution to a standard normally distributed random variable. Let f be a continuously differentiable function. Use Skorohod representation to show that $\sqrt{n}(f(\bar{X}) - f(\mu))/(f'(\mu)\sigma)$ also converges in distribution to a standard normally distributed random variable.

Solution: By Skorohod's representation, there exists random variables Y_n and Y such that

$$Y_n \stackrel{d}{=} \sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right), \quad Y \stackrel{d}{=} Z, \quad \text{and } Y_n \stackrel{a.s.}{\to} Y,$$

where $Z \sim N(0, 1)$. Then, notice

$$\begin{aligned} \sqrt{n}\left(\frac{f(\bar{X}) - f(\mu)}{f'(\mu)\sigma}\right) \stackrel{d}{=} \sqrt{n}\left(\frac{f(\mu + \sigma Y_n/\sqrt{n}) - f(\mu)}{f'(\mu)\sigma}\right) &\stackrel{d}{=} \frac{f(\mu + \sigma Y_n/\sqrt{n}) - f(\mu)}{Y_n \sigma/\sqrt{n}} \cdot \frac{Y_n \sigma}{f'(\mu)\sigma} \\ &\stackrel{a.s.}{\to} f'(\mu) \cdot \frac{Y\sigma}{f'(\mu)\sigma} = Y \stackrel{d}{=} Z \end{aligned}$$

since $Y_n \sigma/\sqrt{n} \stackrel{a.s.}{\to} 0$ by the strong law of large numbers. This implies we have convergence in distribution, which completes the proof.

Problem 3

In this problem, we develop Stirling's formula.

(a) Use integration by parts to show

$$\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-u^2/2}du \sim \frac{1}{\sqrt{2\pi}}\frac{1}{x}e^{-x^2/2}.$$

Solution: Using integration by parts, we have

$$\begin{aligned}\int_x^\infty e^{-u^2/2}du &= \int_x^\infty \frac{u}{u}e^{-u^2/2}du = -\frac{1}{u}e^{-u^2/2}\Big|_x^\infty - \int_x^\infty \frac{1}{u^2}e^{-u^2/2}du \\ &= \frac{1}{x}e^{-x^2/2} - \int_x^\infty \frac{1}{u^2}e^{-u^2/2}du.\end{aligned}$$

Noting that for large x, $\int_x^\infty \frac{1}{u^2}e^{-u^2/2}du \approx 0$, we conclude the result.

(b) Suppose $X_1, X_2, ...$ are iid having expectation 0 and variance 1 and suppose that $a_n \to \infty$. Use the central limit theorem and part (a) to show that

$$P(S_n \geq a_n\sqrt{n}) \sim \frac{1}{\sqrt{2\pi}}\frac{1}{a_n}e^{-a_n^2/2} = e^{-a_n^2(1+\epsilon_n)/2},$$

where $\epsilon_n \to 0$ if $a_n \to \infty$.

Solution: By the central limit theorem, we know that $\sqrt{n}\bar{X}$ converges in distribution to a standard normal random variable, i.e. $S_n/\sqrt{n} \to Z$, where $Z \sim N(0,1)$. Then, for large n,

$$P(S_n \geq a_n\sqrt{n}) = P(S_n/\sqrt{n} \geq a_n) \sim \frac{1}{\sqrt{2\pi}}\int_{a_n}^\infty e^{-u^2/2}du \sim \frac{1}{\sqrt{2\pi}}\frac{1}{a_n}e^{-a_n^2/2}.$$

(c) Suppose $S_n = X_1 + \cdots + X_n$ where the X_i's are independent and each has a Poisson distribution with expectation 1.

(i) Show

$$E\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right] = e^{-n}\sum_{k=0}^n \left(\frac{n-k}{\sqrt{n}}\right)\frac{n^k}{k!} = \frac{n^{n+(1/2)}e^{-n}}{n!}.$$

Solution: Note that $S_n \sim \text{Poisson}(n)$. Then, by law of total expectations,

$$\begin{aligned}E\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right] &= E\left[-\min\left(\frac{S_n-n}{\sqrt{n}}, 0\right)\right] \\ &= \sum_{k=0}^n E\left[-\min\left(\frac{S_n-n}{\sqrt{n}}, 0\right)\Big|S_n = k\right]P(S_n = k) \\ &= \sum_{k=0}^n -\frac{k-n}{\sqrt{n}}\cdot\frac{n^k e^{-n}}{k!} = e^{-n}\sum_{k=0}^n \frac{n-k}{\sqrt{n}}\cdot\frac{n^k}{k!}.\end{aligned}$$

By induction, we can easily show that $e^{-n}\sum_{k=0}^n \left(\frac{n-k}{\sqrt{n}}\right)\frac{n^k}{k!} = \frac{n^{n+(1/2)}e^{-n}}{n!}$.

(ii) Show

$$\left(\frac{S_n - n}{\sqrt{n}}\right)^- \xrightarrow{d} N^-.$$

Solution: We know by the central limit theorem that $\frac{S_n-n}{\sqrt{n}} \xrightarrow{d} N(0,1)$. Now, since $g(x) = -\min(x,0)$ is a continuous function, we have the result.

(iii) Show

$$E\left[\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right] \to E[N^-] = \frac{1}{\sqrt{2\pi}}.$$

Solution: Define $X_n = \left(\frac{S_n-n}{\sqrt{n}}\right)^-$. Since $X_n \xrightarrow{d} N^-$, by Skorohod's representation theorem, there exists random variables $\widetilde{X}_n = \left(\frac{\widetilde{S}_n-n}{\sqrt{n}}\right)^-$ and $\widetilde{N}^-$ such that $\widetilde{X}_n \stackrel{d}{=} X_n$, $\widetilde{N}^- \stackrel{d}{=} N^-$, and $\widetilde{X}_n \xrightarrow{a.s.} \widetilde{N}^-$. Also, see that

$$E\big[\widetilde{X}_n^2\big] = \frac{1}{n}E\big[(\widetilde{S}_n - n)^2\big] = 1.$$

Therefore, we have

$$\int_{\{\widetilde{X}_n \geq c\}} \widetilde{X}_n dP \leq \frac{1}{c}\int_{\{\widetilde{X}_n \geq c\}} \widetilde{X}_n^2 dP \leq \frac{1}{c}E\big[\widetilde{X}_n^2\big] \leq \frac{1}{c} \to 0$$

as $c \to \infty$. This implies we have uniform integrability and hence may interchange limits with expectations. Finally,

$$\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} E\big[\widetilde{X}_n\big] = E\big[\lim_{n\to\infty} \widetilde{X}_n\big] = E\big[\widetilde{N}^-\big] = E[N^-]$$

which proves the result.

(iv) Show Stirling's formula, $n! \sim \sqrt{2\pi}n^{n+(1/2)}e^{-n}$.

Solution: This follows immediately from part (i) and (iii).

Problem 4

Let X and Y be independent Bernoulli random variables on a probability space $(\Omega, \mathcal{B}, P)$ with $X \stackrel{d}{=} Y$ and $P(X=0) = P(X=1) = 0.5$. Let $X_n = Y$ for $n \geq 1$. Show $X_n \xrightarrow{d} X$, but that X_n does NOT converge in probability to X.

Solution: First see that the distribution function for each X_n is

$$F_n(x) = P(X_n \leq x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases},$$

which is the same as the distribution function for X. Thus, $\lim_{n\to\infty} F_n(x) = F(x)$ for all x. However, for $0 < \epsilon < 1$,

$$P(|X_n - X| \geq \epsilon) = P(X_n \neq X) = 1/2$$

for all n. Therefore, X_n does not converge to X in probability.

Problem 5

Suppose $X_1, ..., X_n$ are iid exponentially distributed with mean 1. Let

$$X_{1,n} < \cdots < X_{n,n}$$

be the order statistics. Fix an integer l and show $nX_{l,n} \xrightarrow{d} Y_l$, where $Y_l \sim$ Gamma(l,1). Try doing this (a) in a straightforward way by brute force and then (b) using the Renyi representation (exercise 32 on page 116) for the spacings of order statistics from the exponential density.

Solution: (a) Recall the formula for the lth order statistic is

$$\begin{aligned} f_{X_{l,n}}(x) &= \frac{n!}{(l-1)!(n-l)!}[F_X(x)]^{l-1}[1-F_X(x)]^{n-l}f_X(x) \\ &= \frac{n!}{(l-1)!(n-l)!}[1-e^{-x}]^{l-1}e^{-(n-l)x}e^{-x}. \end{aligned}$$

By the transformation theorem, we have

$$f_{nX_{l,n}}(x) = \frac{(n-1)!}{(l-1)!(n-l)!}[1-e^{-x/n}]^{l-1}e^{-(n-l)x/n}e^{-x/n}.$$

Note that for large n, $1-e^{-x/n} \sim x/n$. Therefore,

$$\begin{aligned} f_{nX_{l,n}}(x) &= \frac{(n-1)!}{(l-1)!(n-l)!}[1-e^{-x/n}]^{l-1}e^{-(n-l)x/n}e^{-x/n} \\ &\sim \frac{(n-1)!}{(l-1)!(n-l)!}[x/n]^{l-1}e^{x} \sim \frac{x^{l-1}e^{-x}}{(l-1)!}\frac{n(n-1)\cdots(n-l+1)}{n^l} \\ &\sim \frac{x^{l-1}e^{-x}}{(l-1)!}, \end{aligned}$$

which is a Gamma$(l,1)$ random variable.
(b) By Renyi representation theorem, we have $Y_k = X_{k,n} - X_{k-1,n} \sim \text{Exp}(n-k+1)$. Note

$$nX_{l,n} = n(Y_l + Y_{l-1} + \cdots + Y_2 + X_{1,n}) = nY_l + nY_{l-1} + \cdots + nY_2 + nX_{1,n}$$

and also $nY_k \sim \text{Exp}\big((n-k+1)/n\big)$ and $nX_{1,n} \sim \text{Exp}(1)$. Observing MFGs of each nY_k,

$$E[e^{nY_k t}] = \frac{(n-k+1)/n}{(n-k+1)/n - t} \to \frac{1}{1-t}$$

which is the MFG of Exp(1). That is, nY_k and $nX_{1,n}$ converge in distribution to Exp(1), and it follows that $nX_{l,n} \xrightarrow{d} \text{Gamma}(l,1)$.