MATH 356 - FALL 2016 - HOMEWORK 5 SOLUTION

This homework set is on **Conditional Expectation, Different Modes of Convergence, Law of Large Number and Central Limit Theorem**. It's due on **Nov. 24th in class**. Late homework will **not** be accepted.

Problem 1. (10pts) Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let X and Y be two independent random variables on Ω with Poisson distributions $P(\lambda_1)$ and $P(\lambda_2)$ respectively ($\lambda_1 > 0$, $\lambda_2 > 0$).

(i) (6pts) For $n = 0, 1, 2, \cdots$, determine the conditional PMF of X given $X + Y = n$, i.e., determine

$$\mathbb{P}(X = k | X + Y = n) \text{ for every } k = 0, 1, \cdots, n.$$

What kind of distribution is this conditional distribution?
Solution: We know that $X + Y$ is a Poisson $P(\lambda_1 + \lambda_2)$ random variable. For every $k = 0, 1, ..., n$,

$$\begin{aligned}
\mathbb{P}(X = k | X + Y = n) &= \frac{\mathbb{P}(X = k, X + Y = n)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = k, Y = n - k)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = k) \cdot \mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} \\
&= \frac{e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}} \\
&= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}.
\end{aligned}$$

So this distribution is a binomial $B\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$ random variable.

(ii) (4pts) Determine $\mathbb{E}[X | X + Y]$ and $\mathbb{E}\left[\frac{1}{X+1} | X + Y\right]$.
Solution: Since the expectation of a binomial $B(n, p)$ is np, $\mathbb{E}[X | X + Y] = \frac{n\lambda_1}{\lambda_1 + \lambda_2}$. By the result of Problem 3 of the midterm Version A (Problem 4 of Version B),

$$\mathbb{E}\left[\frac{1}{X + 1} | X + Y\right] = \frac{1 - \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{n+1}}{\frac{\lambda_1}{\lambda_1 + \lambda_2}(n + 1)} = \frac{(\lambda_1 + \lambda_2)^{n+1} - \lambda_2^{n+1}}{\lambda_1 (\lambda_1 + \lambda_2)^n (n + 1)}.$$

Problem 2. (10pts) Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $(X, Y) : \Omega \to \mathbb{R}^2$ be a random variable (discrete or continuous). Assume that $\mathbb{E}\left[X^2\right] < \infty$, and for every $\omega \in \Omega$, the conditional expectations $\mathbb{E}\left[X^2|Y\right](\omega)$ and $\mathbb{E}[X|Y](\omega)$ both exist. Then, one can define the *conditional variance of* X *given* Y to be the random variable

$$\mathrm{Var}(X|Y) : \Omega \to \mathbb{R}$$

such that

$$\forall \omega \in \Omega,\ \mathrm{Var}(X|Y)(\omega) = \mathbb{E}\left[X^2|Y\right](\omega) - \left[\mathbb{E}[X|Y](\omega)\right]^2.$$

In other words,

$$\mathrm{Var}(X|Y) = \mathbb{E}\left[X^2|Y\right] - \left(\mathbb{E}[X|Y]\right)^2.$$

Show that

$$\mathrm{Var}(X) = \mathrm{Var}\left(\mathbb{E}[X|Y]\right) + \mathbb{E}\left[\mathrm{Var}(X|Y)\right].$$

Proof:

$$\begin{aligned}
\mathrm{Var}\left(\mathbb{E}[X|Y]\right) + \mathbb{E}\left[\mathrm{Var}(X|Y)\right] &= \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] - \left(\mathbb{E}\left[\mathbb{E}[X|Y]\right]\right)^2 \\
& \qquad \mathbb{E}\left[\mathbb{E}\left[X^2|Y\right]\right] - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] \\
&= \mathbb{E}\left[X^2\right] - \left(\mathbb{E}[X]\right)^2 = \mathrm{Var}(X).
\end{aligned}$$

Problem 3. (10pts) Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $\{X_i : i \in \mathbb{N}\}$ and N be random variables on Ω, where X_i's are i.i.d. (independent and identically distributed) $N\left(m, \sigma^2\right)$ random variables (with $m \in \mathbb{R}$ and $\sigma^2 > 0$) and N is a $P(\lambda)$ (with $\lambda > 0$) random variable. Further assume that X_i's are independent of N. For each $n \geq 1$, set $S_n := \sum_{j=1}^{n} X_j$, and define the random variable $S_N : \Omega \to \mathbb{R}$ by

$$\forall \omega \in \Omega,\ S_N(\omega) = S_n(\omega) \text{ if } N(\omega) = n.$$

Then, recall from the example discussed in class that

$$\mathbb{E}[S_N|N] = mN.$$

(i) (5pts) Use the definition of conditional variance provided in Problem 2 to compute $\mathrm{Var}(S_N|N)$.

Solution: As we have already seen in class that, the conditional distribution of S_N given $N = n$ is a Gaussian $N\left(nm, n\sigma^2\right)$ random variable, so

$$\mathbb{E}\left[S_N^2|N\right] = (Nm)^2 + N\sigma^2.$$

Therefore,

$$\begin{aligned}\operatorname{Var}(S_N|N) &= \mathbb{E}\left[S_N^2|N\right] - (\mathbb{E}[S_N|N])^2 \\ &= (Nm)^2 + N\sigma^2 - (Nm)^2 \\ &= N\sigma^2.\end{aligned}$$

(ii) (5pts) Use the result of Problem 2 to compute $\operatorname{Var}(S_N)$.
Solution:

$$\begin{aligned}\operatorname{Var}(S_N) &= \operatorname{Var}(\mathbb{E}[S_N|N]) + \mathbb{E}[\operatorname{Var}(S_N|N)] \\ &= \operatorname{Var}(mN) + \mathbb{E}\left[N\sigma^2\right] \\ &= m^2\operatorname{Var}(N) + \sigma^2\mathbb{E}[N] \\ &= m^2\lambda + \sigma^2\lambda\end{aligned}$$

Problem 4. (10pts) Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $\{X_n : n \in \mathbb{N}\}$ be a sequence of i.i.d. Gaussian $N(0,1)$ random variables on Ω. Denote $S_n := \sum_{j=1}^n X_j$ for each $n \geq 1$.

(i) (4pts) Set $T_n := \frac{S_n}{n}$. Determine, for each $x \in \mathbb{R}$, $F_{T_n}(x)$, the DF of T_n. Does $\lim_{n\to\infty} F_{T_n}(x)$ exists? If so, is the limit a DF? If so, to what distribution does it belong?
Solution: T_n has $N\left(0, \frac{1}{n}\right)$ distribution, so

$$\begin{aligned}F_{T_n}(x) &= \frac{\sqrt{n}}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{nt^2}{2}}dt \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{n}x} e^{-\frac{s^2}{2}}ds \\ &\to \begin{cases}1, & \text{if } x > 0, \\ \frac{1}{2}, & \text{if } x = 0, \\ 0, & \text{if } x < 0.\end{cases}\end{aligned}$$

The limit itself is not a DF. But if X is the Dirac random variable at 0, then

$$\lim_{n\to\infty} F_{T_n}(x) = F_X(x) \quad \forall \text{ continuous point } x \text{ of } F_X.$$

So, $T_n \to X$ in distribution.

(ii) (3pts) Set $R_n := \frac{S_n}{\sqrt{n}}$. Determine, for each $x \in \mathbb{R}$, $F_{R_n}(x)$, the DF of R_n. Does $\lim_{n\to\infty} F_{R_n}(x)$ exists? If so, is the limit a DF? If so, to what distribution does it belong?
Solution: For each $n \in \mathbb{N}$, R_n is an $N(0,1)$ random variable. So the limit of $F_{R_n}(x)$ certainly exists and the limit is the DF of the $N(0,1)$ distribution.

(iii) (3pts) Determine the MGFs of T_n and R_n, i.e, $M_{T_n}(s)$ and $M_{R_n}(s)$ whenever they exist, and study $\lim_{n\to\infty} M_{T_n}(s)$ and $\lim_{n\to\infty} M_{R_n}$. Do they lead to the same conclusion as you found in (i) and (ii)?

Solution: For every $s \in \mathbb{R}$, $M_{T_n}(s) = e^{\frac{s^2}{2n}} \to 1$, and the MGF of X the Dirac random variable at 0 is constantly 1, so $T_n \to X$ in distribution. For every $s \in \mathbb{R}$ and every $n \in \mathbb{N}$, $M_{R_n}(s) = e^{\frac{s^2}{2}}$, which again certainly implies that R_n converges to an $N(0,1)$ random variable in distribution.

Problem 5. **(10pts)** Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $\{X_n : n \in \mathbb{N}\}$ be a sequence of random variables on Ω. Assume that $X_n \to c$ in distribution for some constant $c \in \mathbb{R}$. Show that $X_n \to c$ in probability.

Proof: $X_n \to c$ in distribution, so $\forall x \in \mathbb{R} \setminus \{c\}$,

$$F_{X_n}(x) \to \begin{cases} 0, & \text{if } x < c, \\ 1, & \text{if } x > c. \end{cases}$$

For every $\epsilon > 0$,

$$\begin{aligned} \mathbb{P}(|X_n - c| > \epsilon) &= \mathbb{P}(X_n > c + \epsilon) + \mathbb{P}(X < c - \epsilon) \\ &\leq 1 - F_{X_n}(c+\epsilon) + F_{X_n}(c - \epsilon) \\ &\to 1 - 1 + 0 = 0 \text{ as } n \to \infty. \end{aligned}$$

Problem 6. **(10pts)** Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $\{X_n : n \in \mathbb{N}\}$ be a sequence of random variables on Ω. Assume that $X_n \to c$ in probability for some constant $c \in \mathbb{R}$. Let $g : \mathbb{R} \to \mathbb{R}$ be any continuous and bounded function.

(i) (5pts) Show that $g(X_n) \to g(c)$ in probability.

Proof: Fix arbitrary $\epsilon > 0$, because g is continuous at c, $\exists \delta > 0$, such that

$$|g(y) - g(c)| \leq \epsilon \text{ whenever } |y - c| \leq \delta.$$

Therefore, if $|g(X_n) - g(c)| > \epsilon$, then it must be that $|X_n - c| > \delta$. Hence,

$$\mathbb{P}(|g(X_n) - g(c)| > \epsilon) \leq \mathbb{P}(|X_n - c| > \delta) \to 0.$$

So $X_n \to c$ in probability.

Remark: As one can see from this proof, in order for the conclusion to hold, one only needs g to be continuous at c (no need for g being continuous everywhere and bounded).

(ii) (5pts) Show that $\lim_{n\to\infty}\mathbb{E}\left[g\left(X_n\right)\right]=g\left(c\right)$.

Proof: Since g is bounded, $\sup_{x\in\mathbb{R}}\left|g\left(x\right)\right|:=A<\infty$. For arbitrary $\epsilon>0$, since $g\left(X_n\right)\to g\left(c\right)$ in probability as shown in (i), when n is sufficiently large, $\mathbb{P}\left(\left|g\left(X_n\right)-g\left(c\right)\right|>\epsilon\right)\leq\frac{\epsilon}{2A}$. Therefore,

$$\begin{aligned}\left|\mathbb{E}\left[g\left(X_n\right)-g\left(c\right)\right]\right| &\leq \mathbb{E}\left[\left|g\left(X_n\right)-g\left(c\right)\right|\mathbb{I}_{\{\left|g\left(X_n\right)-g\left(c\right)\right|\leq\epsilon\}}\right]+\mathbb{E}\left[\left|g\left(X_n\right)-g\left(c\right)\right|\mathbb{I}_{\{\left|g\left(X_n\right)-g\left(c\right)\right|>\epsilon\}}\right]\\ &\leq \epsilon+2A\mathbb{P}\left(\left|g\left(X_n\right)-g\left(c\right)\right|>\epsilon\right)\\ &\leq \epsilon+2A\frac{\epsilon}{2A}=2\epsilon.\end{aligned}$$

This is sufficient to imply that $\mathbb{E}\left[g\left(X_n\right)\right]$ tends to $g\left(c\right)$ as $n\to\infty$.

Problem 7. **(10pts)** Given a probability space $(\Omega,\mathcal{S},\mathbb{P})$, let $\{X_n:n\in\mathbb{N}\}$ be a sequence of i.i.d. random variables on Ω with the common distribution being the uniform distribution on $(0,1)$. Set

$$Z_n:=\left(\prod_{i=1}^{n}X_i\right)^{1/n}\quad\text{for each } n\in\mathbb{N}.$$

Show that $Z_n\to a$ in probability for some constant $a\in\mathbb{R}$ and determine the value of a.

Proof: Set $Y_i:=\ln X_i$ for $i\in\mathbb{N}$. Then, $\{Y_i:i\in\mathbb{N}\}$ is i.i.d. random variables with

$$\mathbb{E}\left[Y_1\right]=\int_0^1\ln x dx=-1.$$

Applying the WLLN to $\{Y_i\}$ leads to

$$\frac{\sum_{i=1}^{n}Y_i}{n}\to-1\text{ in probability.}$$

By the result of Problem 6 (i) (by taking the function g to be $g\left(x\right)=e^x$), we have that

$$Z_n=e^{\frac{\sum_{i=1}^{n}Y_i}{n}}\to e^{-1}\text{ in probability.}$$

So $a=1/e$.

Problem 8. **(10pts)** Given a probability space $(\Omega,\mathcal{S},\mathbb{P})$, let $\{X_n:n\in\mathbb{N}\}$ be a sequence of i.i.d. random variables on Ω with finite second moment, i.e., $\mathbb{E}\left[X_1^2\right]=a^2\in\mathbb{R}$. Assume that $\mathbb{E}\left[X_1\right]=m$. For each $n\geq1$, set

$$T_n:=\frac{2}{n\left(n+1\right)}\sum_{i=1}^{n}iX_i.$$

Show that $T_n \to m$ in probability.

Proof: First notice that

$$\mathbb{E}\left[T_n\right] = \frac{2}{n\left(n+1\right)} \sum_{i=1}^{n} im = m.$$

$$\begin{aligned} \mathrm{Var}\left(T_n\right) &= \left(\frac{2}{n\left(n+1\right)}\right)^2 \sum_{i=1}^{n} i^2 \mathrm{Var}\left(X_1\right) \\ &= \frac{4\sum_{i=1}^{n} i^2}{n^2\left(n+1\right)^2}\left(a^2 - m^2\right). \end{aligned}$$

So, $\forall \epsilon > 0$, by Chebyshev's inequality,

$$\mathbb{P}\left(\left|T_n - m\right| > \epsilon\right) \leq \frac{\mathrm{Var}\left(T_n\right)}{\epsilon^2} = \frac{4\left(a^2 - m^2\right)}{\epsilon^2} \cdot \frac{\sum_{i=1}^{n} i^2}{n^2\left(n+1\right)^2} = O\left(\frac{1}{n}\right)$$

which leads to that $T_n \to m$ in probability.

Problem 9. **(10pts)** Given a probability space $(\Omega, \mathcal{S}, \mathbb{P})$, let $\{X_n : n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables on Ω with $\mathbb{E}\left[X_1\right] = m \in \mathbb{R}$ and $\mathrm{Var}\left(X_1\right) = \sigma^2 > 0$, and $\{Y_n : n \in \mathbb{N}\}$ be another sequence of i.i.d. random variables on Ω with $\mathbb{E}\left[Y_1\right] = \mu \in \mathbb{R}$ ($\mu \neq 0$) and $\mathrm{Var}\left(Y_1\right) = \gamma^2 > 0$. For each $n \geq 1$, set $\overline{S_n} := \frac{1}{n}\sum_{i=1}^{n} X_i$, and $\overline{T_n} := \frac{1}{n}\sum_{j=1}^{n} Y_j$, and finally

$$Z_n := \frac{\sqrt{n}\left(\overline{S_n} - m\right)}{\overline{T_n}}.$$

Apply the WLLN and the CLT (Central Limit Theorem) to show that $Z_n \to Z$ in distribution for some random variable Z, and determine the distribution of Z. **Proof:** First, by the WLLN, we know that

$$\overline{T_n} \to \mu \text{ in probability}$$

and hence

$$\frac{1}{\overline{T_n}} \to \frac{1}{\mu} \text{ in probability}$$

On the other hand, by the CLT,

$$\frac{\sqrt{n}\left(\overline{S_n} - m\right)}{\sigma} = \frac{S_n - mn}{\sqrt{n}\sigma} \to Y \text{ in distribution}$$

where Y is an $N\left(0,1\right)$ random variable. So, by Slutsky's theorem,

$$\frac{\sqrt{n}\left(\overline{S_n} - m\right)}{\sigma}\frac{1}{\overline{T_n}} \to \frac{Y}{\mu} \text{ in distribution.}$$

Finally,

$$Z_n = \sigma\frac{\sqrt{n}\left(\overline{S_n} - m\right)}{\sigma}\frac{1}{\overline{T_n}} \to Z = \frac{\sigma}{\mu}Y \text{ in distribution}$$

where Z is an $N\left(0, \frac{\sigma^2}{\mu^2}\right)$ random variable.

Problem 10. **(10pts)** Apply the CLT to i.i.d. Poisson random variables to show that

$$\lim_{n\to\infty} e^{-nt}\sum_{k=0}^{n}\frac{(nt)^k}{k!} = \begin{cases} 0, & \text{if } t > 1, \\ \frac{1}{2}, & \text{if } t = 1, \\ 1, & \text{if } 0 < t < 1. \end{cases}$$

Proof: Let $t > 0$ be fixed. Let $\{X_i : i \in \mathbb{N}\}$ be i.i.d. Poisson $P(t)$ random variables, $S_n = \sum_{i=1}^{n} X_i$. Then S_n is a $P(nt)$ random variable and hence

$$e^{-nt}\sum_{k=0}^{n}\frac{(nt)^k}{k!} = \mathbb{P}(S_n \le n).$$

By the CLT, for every $x \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{S_n - nt}{\sqrt{nt}} \le x\right) \to \Phi(x)$$

where Φ the DF of the $N(0,1)$ distribution. Therefore, when n is sufficiently large,

$$\mathbb{P}(S_n \le n) = \mathbb{P}\left(\frac{S_n - nt}{\sqrt{nt}} \le \frac{n(1-t)}{\sqrt{nt}}\right) \approx \Phi\left(\frac{n(1-t)}{\sqrt{nt}}\right) = \Phi\left(\frac{\sqrt{n}(1-t)}{\sqrt{t}}\right).$$

When $t > 1$, $\sqrt{n}(1-t) \to -\infty$ and hence $\Phi\left(\frac{\sqrt{n}(1-t)}{\sqrt{t}}\right) \to 0$; when $t = 1$, $\Phi\left(\frac{\sqrt{n}(1-t)}{\sqrt{t}}\right) = \frac{1}{2}$ for all n; when $0 < t < 1$, $\sqrt{n}(1-t) \to +\infty$ and hence $\Phi\left(\frac{\sqrt{n}(1-t)}{\sqrt{t}}\right) \to 1$. Thus we have proven the desired conclusion.