E1 222 Stochastic Models and Applications
Hints for Problem Sheet 4.1

1. Given $P[X_n = 0] = 1 - n^{-2}$, $P[X_n = e^n] = n^{-2}$. Show that X_n converge almost surely but not in r^{th} mean.

Hint: An obvious candidate for limit is zero. Define events $A_k^\epsilon = [|X_k - 0| > \epsilon]$ By definition, $X_n \xrightarrow{a.s.} 0$ iff $P[\limsup A_k^\epsilon] = 0, \ \forall \epsilon > 0$. For this we can use Borel-Cantelli lemma. For all $\epsilon > 0$, we have $P[A_k^\epsilon] = P[X_k \neq 0] = k^{-2}$. Hence $\sum_k P[A_k^\epsilon] < \infty$ and hence $P[\limsup A_k^\epsilon] = 0$ thus showing convergence with probability one.

For r^{th} mean convergence, $E[|X_k|^r] = e^{kr} k^{-2}$ which goes to ∞ as $k \to \infty$ and hence the sequence does not converge in r^{th} mean.

2. Given $P[X_n = 0] = 1 - 1/n$, $P[X_n = n^{1/2r}] = 1/n$, X_n are independent. Show that $E|X_n|^r \to 0$ but the sequence does not converge to zero almost surely.

Hint: $E|X_n|^r = \left(n^{1/2r}\right)^r \frac{1}{n} = \frac{1}{\sqrt{n}} \to 0$, as $n \to \infty$.

Taking $A_n^\epsilon = [|X_n - 0| > \epsilon]$, we have $P[A_n^\epsilon] = \frac{1}{n}$. Hence, $\sum_k P[A_k^\epsilon] = \infty$. Since X_n are given to be independent, by Borel Cantelli lemma, $P[\limsup A_k^\epsilon] = 1$ thus showing that the sequence does not converge almost surely.

3. Let $\Omega = [0, \ 1]$ and let P be the usual length measure. Let $X_n = n^{0.25} I_{[0, \, 1/n]}$, $n = 1, 2, \cdots$, where I_A denotes indicator of event A. Does the sequence converge in (i) probability, (ii) r^{th} mean for some r?

Hint: It is easy to see that the distribution of X_n is: $P[X_n = n^{0.25}] = 1/n$ and $P[X_n = 0] = 1 - 1/n$. Hence obvious candidate for limit is zero.

$$P[|X_n - 0| > \epsilon] = 1/n \ \to 0, \quad \text{as} \ \ n \to \infty$$

and hence it converges in probability to zero. We have $E[|X_n|^r = n^{0.25r-1}$. It goes to zero if $r < 4$.

4. Let $X_1, X_2, \cdots$, be random variables with distributions

$$\begin{aligned} F_{X_n}(x) \quad &= 0 \qquad \text{if} \quad x < -n \\ &= \tfrac{x+n}{2n} \quad \text{if} \quad -n \leq x \leq n \\ &= 1 \qquad \text{if} \quad x \geq n \end{aligned}$$

Does $\{X_n\}$ converge in distribution?

Hint: It is easy to see that X_n is uniform over $[-n,\ n]$. Since we cannot have a uniform density over the entire real line, intuitively, we do not expect the sequence to converge in distribution.

Here, the sequence of functions F_{X_n} converges pointwise to the constant function 0.5. That is, $\lim_{n\to\infty} F_{X_n}(x) = 0.5, \forall x$. The limit function is not a distribution function and hence the X_n does not converge in distribution. You can get an idea of the limit function by plotting F_{X_n} for a couple of large n values.

To show that F_{X_n} converges to the constant function you can proceed as follows. Fix x. Now, given any $\epsilon > 0$ we should show that there is an N such that $|F_{X_n}(x) - 0.5| < \epsilon$ if $n > N$. Once again, I hope you can see this by looking at the expression or graph of the function. Once we fix an x, if n is sufficiently large, x would be in the interval $[-n,\ n]$. Now, if n is sufficiently large (remembering that x is fixed) we can see that $\frac{x+n}{2n}$ is approximately half.

5. Consider a Probability space $(\Omega, \mathcal{F}, P)$ where $\Omega = \{1, 2, \cdots\}$, $\mathcal{F}$ is the power set of Ω and $P(\{i\}) = q_i,\ \forall i$. Note that we would have $q_i \geq 0, \forall i$ and $\sum_i q_i = 1$. Let $X_1, X_2, \cdots$ be a sequence of discrete random variables defined on this space given by

$$\begin{aligned} X_n(\omega) &= 1 \text{ if } n \leq \omega \\ &= 0 \text{ otherwise} \end{aligned}$$

Does the sequence converge in (i) Probability, (ii) almost surely.

Hint: We have

$$\{\omega \in \Omega : X_n(\omega) = 1\} = \{\omega \in \Omega : \omega \geq n\} = \{n, n+1, \cdots\}$$

Hence, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P[|X_n - 0| > \epsilon] = \lim_{n\to\infty} P[X_n = 1] = \lim_{n\to\infty} \sum_{k=n}^{\infty} q_k = 0$$

because $\sum_i q_i = 1$. Hence X_n converges in probability to zero.

To show almost sure convergence, Borel-Cantelli lemma is not useful here. In general we may not have $\sum_{n=1}^{\infty} \sum_{k=n}^{\infty} q_k < \infty$.

By definition of $X_n(\omega)$, at any ω, $X_n(\omega) = 0, \forall n > \omega$. Hence $X_n(\omega) \to 0, \forall \omega$ and hence we have almost sure convergence.

6. Let $X_1, X_2, \cdots$ be iid Gaussian random variables with mean zero and variance unity. Let $\bar{X}_n = (X_1 + \cdots + X_n)/n$. Let F_n be the distribution function of $\bar{X}_n$. Find Lim F_n. Is this a distribution function?

Hint: By law of large numbers we know $\bar{X}_n$ converges to 0 (the mean). A constant is like a discrete random variables that takes only one value. Hence, the distribution function of the constant 0 is $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x \geq 0$. Since convergence in probability or almost sure convergence implies convergence in distribution, the sequence $F_{\bar{X}_n}$ converges to the df F which is given above

This completes the the answer to the question. But since X_i are iid Gaussian, we know that $\bar{X}_n$ is Gaussian with zero mean and variance $1/n$. Hence we know the sequence of distribution functions. It is instructive to look at this sequence of distribution functions and show that the limit is the function F defined above

7. Let $X_1, X_2, \cdots$ be a sequence of discrete random variables with X_n being uniform over the set $\{\frac{1}{n}, \frac{2}{n}, \cdots, \frac{n}{n}\}$. Does the sequence $\{X_n\}$ converge in distribution?

Hint: The sequence here converges to the uniform distribution over [0, 1] You can see this easily by drawing the functions. The df of X_n is a staircase function that starts at the origin has equal height jumps at $1/n$, $2/n$, and so on and reaches 1. (Note that X_n can take n distinct values and hence each jump would have height of $1/n$). Now draw the df of X_{2n}. Now you would be able to see the sequence converges to df of uniform density. Now prove that the sequence of df's converges point-wise. Fix any x in (0, 1). You want to show that $|F_{X_n}(x) - x| < \epsilon$ if n is large. For this, note that if $\frac{k}{n} \leq x < \frac{k+1}{n}$, then we know that $F_{X_n}(x) = \frac{k}{n}$. Also, for all such x, $|x - k/n| < 1/n$.

8. Let $\{X_n\}$ be a sequence of random variables converging in distribution to a continuous random variable X. Let a_n be a sequence of positive numbers such that $a_n \to \infty$ as $n \to \infty$. Show that X_n/a_n converges to zero in probability.

Hint: We are given $a_n > 0$ and they go to infinity.

$$P\left[\left|\frac{X_n}{a_n} - 0\right| > \epsilon\right] = P\left[|X_n| > a_n\epsilon\right] = F_{X_n}(-a_n\epsilon) - P[X_n = -a_n\epsilon] + 1 - F_{X_n}(a_n\epsilon)$$

So, we essentially need to show $\lim_{n\to\infty} F_{X_n}(a_n\epsilon) = 1$ (and similarly $F_{X_n}(-a_n\epsilon) \to 0$, $P[X_n = -a_n\epsilon] \to 0$).

We know that $\lim_{n\to\infty} F_{X_n}(x) = F(x), \forall x$ where F is the limit distribution function. This is true for all x because we are given that the sequence converges to a continuous rv and hence the limit df F is continuous at all x and hence the pointwise converges holds for all x. We also know the $a_n\epsilon \to \infty$ because we are given a_n goes to infinity. Hence you can conclude $F_{X_n}(a_n) \to F(\infty) = 1$. (Similarly, $F_{X_n}(-a_n\epsilon) \to F(-\infty) = 0$ and so on). This is the expected answer.

However, there is an important issue here. The question is if $F_{X_n}(x) \to F(x), \forall x$ and if $a_n \to \infty$, does $F_{X_n}(a_n) \to F(\infty) = 1$? In general, if $y_n \to y$ and $g_n(x) \to g(x)$, $\forall x$, then, it does not necessarily imply $g_n(y_n) \to g(y)$. For that you need the sequence of functions to converge *uniformly*. But in our case since the sequence is of distribution functions, it can be shown that this holds.

9. Let $X_1, X_2, \cdots$ be independent normally distributed random variables having mean zero and variance σ^2.

   (a). What is the mean and variance of X_1^2?

   (b). How should $P[X_1^2 + X_2^2 + \cdots + X_n^2 \leq x]$ be approximated in terms of standard normal distribution?

   (c). Suppose $\sigma^2 = 1$. Find (approximately) $P[80 \leq X_1^2 + \cdots + X_{100}^2 \leq 120]$.

   (d). Find c such that (approximately) $P[100 - c \leq X_1^2 + \cdots + X_{100}^2 \leq 100 + c] = 0.95$.

Answer: (a). Since $EX_1 = 0$, we have $EX_1^2 = \text{Var}(X_1) = \sigma^2$. To find variance of X_1^2 we need to find EX_1^4. Since X_1 is Gaussian with mean zero and variance σ^2, its moment generating function is given by $M_{X_1}(t) = \exp(0.5t^2\sigma^2)$. By differentiating this four times (which is easy to do if you expand it in Taylor series), we get $EX_1^4 = 3\sigma^4$. Hence, $\text{Var}(X_1^2) = E[(X_1^2)^2] - (E[X_1^2])^2 = 2\sigma^4$.

(b). Let $S_n = \sum_{i=1}^n X_i^2$. Since X_i are iid, $ES_n = n\sigma^2$ and $\mathrm{Var}(S_n) = 2n\sigma^4$. Hence

$$P[S_n \le x] = P\left[\frac{S_n - n\sigma^2}{\sigma^2\sqrt{2n}} \le \frac{x - n\sigma^2}{\sigma^2\sqrt{2n}}\right] \approx \Phi\left(\frac{x - n\sigma^2}{\sigma^2\sqrt{2n}}\right)$$

(c). From the above it is easy to see that

$$P[a \le S_n \le b] \approx \Phi\left(\frac{b - n\sigma^2}{\sigma^2\sqrt{2n}}\right) - \Phi\left(\frac{a - n\sigma^2}{\sigma^2\sqrt{2n}}\right)$$

Here, we are given, $\sigma^2 = 1$, $n = 100$, $a = 80$ and $b = 120$. Hence

$$P[80 \le S_{100} \le 120] \approx \Phi\left((120 - 100)/10\sqrt{2}\right) - \Phi\left((80 - 100)/10\sqrt{2}\right)$$

Hence

$$P[80 \le S_{100} \le 120] \approx 2\Phi(\sqrt{2}) - 1 \approx 2\Phi(1.41) - 1 = 0.84$$

(d). As above, we have

$$P[100 - c \le S_{100} \le 100 + c] \approx 2\Phi\left(c/10\sqrt{2}\right) - 1$$

Equating this to 0.95 we get $c/10\sqrt{2} = \Phi^{-1}(0.975) = 1.96$. Hence $c = 19.6 * \sqrt{2} = 27.7$.

10. Candidates A and B are contesting an election and 55% of the electorate favour B. What is the (approximate) probability that in a sample of size 100 atleast one-half of the people sampled favour A.

Answer: Let X_i be iid random variables with $P[X_i = 1] = 0.45 = 1 - P[X_i = 0]$. Thus, X_i is an indicator of whether i^{th} person sampled favours A. Let $S_{100} = \sum_{i=1}^{100} X_i$. Hence what we want is $P[S_{100} > 50]$. Since S_{100} is integer-valued and we are using a CLT approximation, this probability is often written as $P[S_{100} > 50.5]$. (This is used for better approximation whenever we are approximating binomial distribution with normal distribution). Note that $ES_{100} = 45$ and $\mathrm{Var}(S_{100}) = 100 * 0.45 * (1 - 0.45) = 24.75$ (We have $\sqrt{24.75} = 4.97 \approx 5$).

$$P[S_{100} > 50.5] = P\left[\frac{S_{100} - 45}{0.497} > \frac{50.5 - 45}{4.97}\right] \approx 1 - \Phi(1.11) = 1 - 0.86 = 0.14$$

Comment: In an exam it is alright if you use $P[S_{100} > 50]$ instead of $P[S_{100} > 50.5]$

11. A university has 300 vacancies for research students. Since not all students offered admission would accept, the university sends out offers of admission to 400 students. By past experience the university knows that only 70% of students offered admission would accept the offer. Calculate the approximate probability that more than 300 students would accept the offer of admission.

Hint: Let X_i be iid random variables with $P[X_i = 1] = 0.7 = 1 - P[X_i = 0]$. Let $S_{400} = \sum_{i=1}^{400} X_i$. The probability we want is $P[S_{400} > 300]$. We know $EX_i = 0.7$ and $\mathrm{Var}(X_i) = 0.21$. Now I hope you can solve the problem.

12. A fair coin is tossed until 100 heads appear. Find (approximately) the probability that atleast 230 tosses will be necessary.

Answer: The event of at least 230 tosses being needed is same as the event that in 229 tosses the number of heads is less than or equal to 99. Let X_i be iid random variables taking values 0 and 1 with equal probability. Let $S_n = \sum_{i=1}^{n} X_n$. Now, S_n represents the number of heads in n tosses of a fair coin. Note that $EX_i = 0.5$ and $\mathrm{Var}(X_i) = 0.25$. Hence $ES_n = 0.5n$ and $\mathrm{Var}(S_n) = 0.25n$. Hence, the probability we need is

$$P[S_{229} \leq 99] = P\left[\frac{S_{229} - 0.5 * 229}{\sqrt{229 * 0.25}} \leq \frac{99 - 0.5 * 229}{\sqrt{229 * 0.25}}\right]$$

$$\approx \Phi\left(\frac{99 - 0.5 * 229}{\sqrt{229 * 0.25}}\right) = \Phi(-2.05) = 0.02$$