1. (Durrett 2.1.5) Let D be the set of pointmasses of μ

$$\int\int 1_{x+y=0}d\mu d\nu = \int\int 1_{x=-y}d\mu d\nu = \int \mu\{-y\}\,d\nu = \sum_{-y\in D}\mu\{-y\}\,\nu y = \sum_y \mu\{-y\}\,\nu y$$

$\square$

2. (Durrett 2.1.8) $X_1,\ldots,X_4$ independent taking values ± 1 with probability $\frac{1}{2}$. $Y_1 = X_1 X_2$, $Y_2 = X_1 X_3$, $Y_3 = X_2 X_4$, $Y_4 = X_3 X_4$, by symmetry and the fact that $\{Y_i = 1\}$ is the desired π-system, it suffices to check that $P(Y_1 = Y_2 = Y_3 = 1) = \frac{1}{8}$, but this is striaghtforward by using the definition of the X_i. $\square$

3. (Durrett 2.1.12) The combinatorial proof of $\sum_j \binom{n}{j}\binom{m}{k-j} = \binom{n+m}{j}$ writes itself. This is a straight-forward computation

$$P(X+Y=k) = \sum_j P(X=j)P(Y=k-j) = \sum_j\sum_j \binom{n}{j}\binom{m}{k-j}p^{j+k-j}(1-p)^{n+m-k+j-j}$$

$$= p^k(1-p)^{n+m-k}\sum_j \binom{n}{j}\binom{m}{k-j} = p^k(1-p)^{n+m-k}\binom{n+m}{j}$$

$\square$

4. (Durrett 2.5.3) Define $Y_n = \frac{\sin(n\pi t)}{n}X_n$, then since $X_n \sim N(0,1)$ we get $\operatorname{Var} Y_n = \frac{|\sin(n\pi t)|}{n^2} \le 1/n^2$. Then since $\sum_1^\infty \operatorname{Var} Y_n < \infty$ and $Y_1, Y_2, \ldots$ are independent, we get from the consequence of Kolmogorov maximal inequality that $\sum_1^\infty Y_n$ converges almost surely. $\square$

5. (Durrett 2.5.6) We can use the Kolmogorov 3-series test. First notice,

$$E[\psi(X_n)] = E\left[|X_n|1_{|X_n|>1} + X_n^2 1_{|X_n|\le 1}\right]$$

So in particular by the comparison test,

$$\sum_1^\infty E\left[|X_n|1_{|X_n|>1}\right] < \infty \qquad\qquad \sum_1^\infty E\left[X_n^2 1_{|X_n|\le 1}\right] < \infty$$

The latter is the series $\sum_1^\infty \operatorname{Var}(X_n 1_{|X_n|\le 1})$, another of the series is relatively free, namely

$$\sum_1^\infty P(|X_n| > 1) \le \sum E\left[|X_n|1_{|X_n|>1}\right] < \infty$$

Finally, we have $\left|E[X_n 1_{|X_n|>1}]\right| \le E\left[|X_n|1_{|X_n|>1}\right]$, so that $\sum_1^\infty E\left[X_n 1_{|X_n|>1}\right]$ converges, since

$$0 = \sum_1^N E[X_n] = \sum_1^N E\left[X_n 1_{|X_n|>1} + X_n 1_{|X_n|\le 1|}\right] = \sum_1^N E\left[X_n 1_{|X_n|>1}\right] + \sum_1^N E\left[X_n 1_{|X_n|\le 1|}\right]$$

we have

$$\sum_1^\infty E\left[X_n 1_{|X_n|\le 1|}\right] = -\sum_1^\infty E\left[X_n 1_{|X_n|>1}\right] \in \mathbb{R}$$

Thus X_n satisfy the hypothesis of the Kolmogorov 3-series test and converge almost surely. $\square$

6. (Durrett 2.5.8) Write $Y = \log^+ |X_1|$ and assume first that $EY = \infty$, then for any $c > 0$, we find that

$$EY = \int P(Y > t)dt = \sum_1^\infty \int_{(n-1)c}^{nc} P(Y > t)dt \leq \sum_0^\infty cP(Y > nc)$$

which in turn implies that $\sum_1^\infty P(\log^+ |X_n| > nc) = \sum_1^\infty P(Y > nc) = \infty$, therefore by Borel-Cantelli (ii), we find that $P(X_n > e^{nc}$ i.o.$) = 1$, so for any $c \neq 0$ we find that $|X_n| c^n > 1$ i.o. almost-surely. This of course implies that $\sum_1^\infty |X_n| c^n$ has radius of convergence zero almost surely.

Now, conversely suppose $E\log^+ |X_n| < \infty$, and let $0 < c < 1$, then choose $\gamma > 0$ so that $e^\gamma c < 1$ we get another layer cake estimate,

$$EY = \int P(Y > t)dt = \sum_1^\infty \int_{(n-1)\gamma}^{n\gamma} P(Y > t)dt \geq \sum_1^\infty \gamma P(Y > n\gamma)$$

Since $\sum_1^\infty \gamma P(Y > n\gamma) = \sum_1^\infty P(\log^+ |X_n| > n\gamma)$ by Borel Cantelli, we have $|X_n| > e^{n\gamma}$ only finitely many times. Letting N so that $n \geq N$ implies $|X_n| \leq e^{n\gamma}$, we get

$$\sum_1^\infty c^n |X_n| \overset{\text{a.s.}}{\leq} \sum_1^{N-1} c^n |X_n| + \sum_N^\infty (ce^\gamma)^n < \infty$$

So that the series converges a.s. for any $c < 1$. Now letting $c > 1$, we once again use Borel-Cantelli (ii), we first check the hypothesis

$$\sum_1^\infty P(|X_n| c^n > 1) = \sum_1^\infty P(|X_1| > \frac{1}{c^n}) = \infty$$

Since $\lim_{n\to\infty} P(|X_1| > \frac{1}{c^n}) = P(X_1 \neq 0) > 0$ by assumption, it follows that $P(|X_n| c^n > 1$ i.o.$) = 1$, so the series diverges a.s. for $c > 1$. $\qquad\square$

7. (Durrett 2.5.9) First define some notation to make the problem read more easily,

$$A = \left\{ \max_{m < j \leq n} |S_{m,j}| > 2a \right\} \qquad\qquad A_j = \{|S_{m,j}| > 2a \text{ and } |S_{m,i}| \leq 2a \text{ for } i < j|\}$$
$$B_k = \{|S_{k,n}| \leq a\}$$

Now we get that $A = \bigsqcup_{m+1}^n A_j$, moreover A_j, B_j are independent since $A_j \in \mathcal{F}(X_m, \dots, X_j)$ and $B_j \in \mathcal{F}(X_{j+1}, \dots, X_n)$, finally notice that for each j, we have $S_{m,n} = S_{m,j} + S_{j,n}$ so that $|S_{m,n}| \geq |S_{m,j}| - |S_{j,n}|$ which namely implies that

$$\{|S_{m,n}| > a\} \supset \bigsqcup_{m+1}^n A_j \cap B_j$$

At this point we are basically done, we only need to check a simple computation

$$P(|S_{m,n}| > a) \geq P(\bigsqcup_{m+1}^n (A_j \cap B_j)) = \sum_{m+1}^n P(A_j \cap B_j) = \sum_{m+1}^n P(A_j)P(B_j)$$
$$\geq \sum_{m+1}^n P(A_j) \min_k P(B_k) = P(A) \min_k P(B_k)$$

$\qquad\square$

8. (Durrett 2.5.10) This problem uses a few techniques I felt may be useful, the first of which is a set of nonconvergence can be written nicely as a limsup of sets. The second technique is to show the measure of the limsup of a set is zero is not always possible using Borel Cantelli, but in many cases we may have descending sequences of sets, in which case we can simply apply continuity from above.

Letting $1 \leq M < n < \infty$, we get the inequality

$$\max_{M \leq k < m \leq n} |S_{k,m}| = \max_{M \leq k < m \leq n} |S_{k,n} - S_{m,n}| \leq 2 \max_{M \leq m \leq n} |S_{m,n}|$$

Now we can use this to rewrite the inequality from exercise (2.5.9),

$$P(|S_{M,n}| > \epsilon) \geq P(\max_{M \leq k < m \leq n} |S_{k,m}| > 4\epsilon) \min_{M < k \leq n} P(|S_{k,n}| \leq \epsilon)$$

Taking the limit as $n \to \infty$ we get

$$P(|S_{M,\infty}| > \epsilon) \geq P(\sup_{k,m \geq M} |S_{k,m}| > 4\epsilon) \inf_{k > M} P(|S_{k,\infty}| < \epsilon)$$

Since S_n is cauchy in measure, the term on the left goes to zero and the infimum term goes to 1 as $M \to \infty$, it follows that $\lim_{M \to \infty} P(W_M^\epsilon) = 0$ for any positive ϵ where $W_M^\epsilon := \{\sup_{k,m \geq M} |S_{k,m}| > \epsilon\}$. We may notice in our case that $W_1^\epsilon \supset W_2^\epsilon \supset \cdots$, so that

$$P(W_M^\epsilon \text{ i.o.}) = P(\bigcap_1^\infty W_M^\epsilon) = \lim_{M \to \infty} P(W_M^\epsilon) = 0$$

We also have that $\left\{W_M^1 \text{ i.o.}\right\} \subset \left\{W_M^{1/2} \text{ i.o.}\right\} \subset \left\{W_M^{1/3} \text{ i.o.}\right\} \subset \cdots$ so that $P(\bigcup_{n=1}^\infty W_M^{1/n} \text{ i.o.}) = \lim_{n \to \infty} P(W_M^{1/n} \text{ i.o.}) = 0$ it follows that $S_n(x)$ is cauchy (hence convergent) almost everywhere, so that $S_n \to S_\infty$ a.s. because $x \notin \bigcup_{n=1}^\infty \left\{W_M^{1/n} \text{ i.o.}\right\}$ implies $S_n(x)$ is cauchy. $\qquad \square$

9. (Durrett 2.5.11) Let $\epsilon > 0$, then suppose that for some sequence n_p, we have

$$\min_{k \leq n_p} P\left(\left|\sum_k^{n_p} x_j\right| < n_p \epsilon\right) \to 0$$

$k_p := \text{argmin}_{k \leq n_p} P\left(\left|\sum_k^{n_p} x_j\right| < n_p \epsilon\right)$ Then we can write

$$\frac{\left|S_{n_p} - \sum_{k_p}^{n_p} X_j\right|}{n_p} = \frac{|S_{k_p}|}{n_p}$$

Which gives the inequality

$$\frac{|S_{k_p}|}{n_p} \geq \frac{\left|\sum_{k_p}^{n_p} X_j\right|}{n_p} - \frac{|S_{n_p}|}{n_p}$$

So that $\left\{|S_{k_p}|/n_p < \epsilon\right\} \subset \left\{\left|\sum_{k_p}^{n_p} X_j\right|/n_p < 2\epsilon\right\} \cup \left\{|S_{n_p}|/n_p \geq \epsilon\right\}$, taking probabilities we find that

$$P\left(|S_{k_p}|/n_p < \epsilon\right) \leq P\left(\left|\sum_{k_p}^{n_p} X_j\right|/n_p < 2\epsilon\right) + P\left(|S_{n_p}|/n_p \geq \epsilon\right)$$

$P\left(|S_{n_p}|/n_p \geq \epsilon\right) \to 0$ by assumption, and checking the left hand side, we find that along any subsequence with k_p uniformly bounded by some N, we have $|S_{k_p}|/n_p \to 0$ pointwise, since each of $|S_1|/n_p, \ldots, |S_N|/n_p \to 0$ pointwise as $p \to \infty$. Whereas on any increasing subsequence of the k_p, we have

$\left|S_{k_p}\right|/n_p \leq S_{k_p}/k_p \xrightarrow{P} 0$, so that every subsequence of k_p has a further subsequence converging to zero in probability, it follows that $P\left(\left|S_{k_p}\right|/n_p < \epsilon\right) \to 1$, which in turn implies that $P\left(\left|\sum_{k_p}^{n_p} X_j\right|/n_p < 2\epsilon\right) \to 1$ which contradicts our previous assumption. Hence we find that $\liminf_n \min_{k \leq n} P\left(\left|\sum_k^n x_j\right| < n\epsilon\right) > 0$. Now we apply the inequality of exercise (2.5.9) to notice that

$$P\left(\max_{1 \leq k \leq n} |S_k|/n > 2\epsilon\right) \min_{1 \leq j \leq n} P\left(\left|\sum_j^n X_j\right|/n \leq \epsilon\right) \leq P\left(|S_n|/n > \epsilon\right)$$

Then since the right hand side goes to zero by assumption, and $P\left(\left|\sum_j^n X_j\right|/n \leq \epsilon\right)$ is bound away from zero for sufficiently large n we must have $P\left(\max_{1 \leq k \leq n} |S_k|/n > 2\epsilon\right) \to 0$ as desired. $\square$

10. (Durrett 3.2.9) We can use a mesh to show that $||F - F_n||_u = 0$. Namely, let $\epsilon > 0$, then we can choose x_0 so that $F(x_1) < \epsilon$ and y so that $F(y) > 1 - \epsilon$, then F is uniformly continuous on $[x_0, y]$, so in particular there is some $\delta > 0$ such that $|w - z| < \delta \implies |F(w) - F(z)| < \epsilon$ on $[x_0, y]$, then take $x_j = x_0 + j\frac{\delta}{2}$ for $j \in \{1, \ldots, N-1\}$, so that $y - \delta/2 < x_{N-1} \leq y$ and denote $y = x_N$. Then take M_j, so that $n \geq M_j \implies |F_n(x_j) - F(x_j)| < \epsilon$ and define $M = \max_{0 \leq j \leq N} M_j$. It follows that for $n \geq N$, we have for $x \in (-\infty, x_0)$,

$$F_n(x) \leq F_n(x_0) < F(x_0) + \epsilon < 2\epsilon$$

and for $x \in (x_N, \infty)$

$$F_n(x) \geq F_n(x_N) > F(x_N) - \epsilon > 1 - 2\epsilon$$

and on these ranges $F(x) < \epsilon$ and $F(x) > 1 - \epsilon$ respectively, meaning $|F(x) - F_n(x)| < 2\epsilon$. Finally, suppose that $x \in [x_{j-1}, x_j]$ for some j, then we have

$$F(x) - 2\epsilon < F(x_{j-1}) - \epsilon \leq F_n(x_{j-1}) \leq F_n(x) \leq F_n(x_j) \leq F(x_j) + \epsilon < F(x) + 2\epsilon$$

So that $||F - F_n||_u < 2\epsilon$ and since ϵ was arbitrary, $||F - F_n||_u \to 0$. $\square$

11. (Durrett 3.2.12) One direction is much easier, so first suppose that $X_n \implies c$, then for any $\epsilon > 0$, we have $P(X_n \leq c - \epsilon) \to P(c \leq c - \epsilon) = 0$, and $P(X_n > c + \epsilon) = 1 - P(X_n \leq c + \epsilon) \to 1 - P(c \leq c + \epsilon) = 0$, so that
$$P(|X_n - c| > \epsilon) \leq P(X_n > c + \epsilon) + P(X_n \leq c - \epsilon) \to 0$$

For the (more general) converse, I should first mention that the DCT applies to convergence in probability (every subsequence has a subsequence converging almost surely which will have the same DCT limit) we can use the Portmanteau lemma, which gives equivalence to weak-* convergence on bounded continuous functions. Namely, we want to show that $\int f d\mu_n \to \int f d\mu$ for all bounded continuous f. Now letting f be bounded and continuous, we have that $E|f| < \infty$ so we can apply DCT, observe from the change of variables formula,

$$\int f d\mu_n = E[f \circ X_n] \xrightarrow{\text{DCT}} E[f \circ X] = \int f d\mu$$

$\square$

12. (Durrett 3.2.13) - Converging Together Lemma We use the previous problem, $Y_n \implies c$, so that $Y_n \xrightarrow{P.} c$ now we can write $P(X \leq \omega) = P(X + c \leq \omega + c)$, if $\mathcal{C}_X$ is the set of points where F_X is continuous, the set of points where F is continuous is $\mathcal{C}_X + c$. Now let $\epsilon > 0$, and $x \in \mathcal{C}_X + c$, then choose δ such that $|x - c - y| \leq \delta$ implies $|F(x - c) - F(y)| < \epsilon$ and $x - c - \delta$ is a point of continuity for F

$$\{X_n + Y_n \leq x\} \supset \{X_n \leq x - \delta - c\} \cap \{Y \in (c - \delta, c + \delta)\}$$

We can then write

$$P(X_n + Y_n \leq x) \geq F_n(x - c - \delta) - P(Y_n \notin (c - \delta, c + \delta))$$
$$\geq F_X(x - c) - |F_n(x - c - \delta) - F_X(x - c - \delta)| - |F_X(x - c - \delta) - F_X(x - c)| - P(Y_n \notin (c - \delta, c + \delta))$$

Now taking n large enough, we get $|F_n(x - c - \delta) - F_X(x - c - \delta)| < \epsilon$ and $P(Y_n \notin (c - \delta, c + \delta)) = P(|Y_n - c| \geq \delta) < \epsilon$ so that

$$P(X_n + Y_n \leq x) \geq F_X(x - c) - 3\epsilon = F(x - c) - 3\epsilon$$

For the converse inequality,

$$P(X_n + Y_n \leq X) = P(X_n + Y_n \leq x, |Y_n - c| \geq \delta) + P(X_n + Y_n \leq x, |Y_n - c| < \delta)$$
$$\leq P(|Y_n - c| \geq \delta) + P(X_n \leq x - c - \delta) \leq F_X(x - c) + 3\epsilon$$

by the same inequalities as above $\square$

13. (Durrett 3.2.15) We let $Y_j \sim N(0, 1)$, and $\hat{X}_n^j = Y_j(\frac{n}{\sum_1^n Y_j^2})$, we will show later that $\hat{X} \stackrel{\text{dist.}}{=} X$, and that $\hat{X}_n^1 \implies N(0, 1)$. Starting with the latter, Y_j^2 are i.i.d. With $E\left|Y_j^2\right| = EY_j^2 = \text{var}(Y_j) + (EY_j)^2 = 1$, whence we can apply the strong law of large numbers to $\frac{\sum_1^n Y_j^2}{n} \stackrel{\text{a.s.}}{\longrightarrow} 1$ so that $\hat{X}_n^1 \stackrel{\text{a.s.}}{\longrightarrow} Y_1 \sim N(0, 1)$.

Now to see that $\hat{X} \stackrel{\text{dist.}}{=} X$, we will show that $\hat{X}$ is uniformly distributed on the sphere of radius $\sqrt{n}$, first note that it indeed takes values on this sphere since $\sqrt{\sum_1^n \left(Y_j \left(\frac{n}{\sum_1^n Y_k^2}\right)\right)^2} = \sqrt{n}$. Now we can use that the uniform distribution is the unique distribution which is preserved by rotation on the $\sqrt{n}$-sphere. Denoting $Y = (Y_1, \ldots, Y_n)$, Y having distribution F can use characteristic functions to check that

$$\varphi_{AY}(t) = \int e^{it^T AY} dF = \int e^{i(A^T t)^T Y} dF = \varphi_Y(A^T t) = e^{||A^T t||^2/2} = e^{||t||^2/2} = \varphi_Y(t)$$

So that $Y = AY$, now $A\hat{X} = \sqrt{n}\frac{AY}{||Y||} = \sqrt{n}\frac{Y}{||Y||}$. $\square$

14. (Durrett 1.7.5) The following is used in the proof of Fourier inversion, namely $\int_0^\infty \frac{\sin x}{x} = \frac{\pi}{2}$. Moreover the trick $\frac{1}{x} = \int_0^\infty e^{-xy} dy$, then using Fubini is frequently useful for integration. Now on with the proof,

Let $I := \int_0^a e^{-xy} \sin x \, dx$, then

$$I = \left[-\cos x e^{-xy} - y \int \cos x e^{-xy}\right]_{x=0}^a = \left[-\cos x e^{-xy} - y e^{-xy} \sin x - y^2 \int e^{-xy} \sin x\right]_{x=0}^a$$
$$= 1 - \cos a e^{-ay} - y e^{-ay} \sin a + y^2 I$$

So that $I = \frac{1}{1+y^2} - \cos a \frac{e^{-ay}}{1+y^2} - \sin a \frac{y e^{-ay}}{1+y^2}$

Now it is easy to see $\int_0^\infty I \, dy = \pi/2 - \cos a \int_0^\infty \frac{e^{-ay}}{1+y^2} dy - \sin a \int_0^\infty \frac{y e^{-ay}}{1+y^2}$. Since Tolleni's theorem always holds for positive functions, we can check the assumptions of Fubini on $|e^{-xy} \sin x|$, using $|\sin x| \leq |x|$

$$\int_0^a \int_0^\infty \left|e^{-xy} \sin x\right| dy \, dx = \int_0^a \frac{|\sin x|}{x} dx \leq a$$

All of the work has been done, so now we can breeze through the proof

$$\int_0^a \frac{\sin x}{x} = \int_0^a \int_0^\infty e^{-xy} \frac{\sin x}{x} dy \, dx \stackrel{\text{fubini}}{=} \int_0^\infty \int_0^a e^{-xy} \frac{\sin x}{x} dx \, dy$$
$$= \int_0^\infty \frac{1}{1+y^2} - \cos a \frac{e^{-ay}}{1+y^2} - \sin a \frac{y e^{-ay}}{1+y^2}$$
$$= \pi/2 - \cos a \int_0^\infty \frac{e^{-ay}}{1+y^2} dy - \sin a \int_0^\infty \frac{y e^{-ay}}{1+y^2} dy$$

This gives the desired inequality, namely

$$\left| \int_0^a \frac{\sin x}{x} - \pi/2 \right| \le |\cos a| \int_0^\infty e^{-ay} dy + |\sin a| \int_0^\infty y e^{-ay} dy \le \frac{1}{a} + \frac{1}{a^2} \overset{a \ge 1}{\le} \frac{2}{a}$$

$\square$

15. (Durrett 3.3.1) Let φ be the characteristic function of X, then let Y be i.i.d. with $-X$, so that

$$\varphi_{X+Y}(t) = E[e^{it(X+Y)}] \overset{\text{ind.}}{=} E[e^{itX}]E[e^{itY}] = \varphi(t)\overline{\varphi(t)} = |\varphi(t)|^2$$

To get that $\Re\varphi$ is a characteristic function, note first that it must be

$$\Re\varphi(t) = E\left[\frac{1}{2}(e^{itX} + e^{-itX}) \right]$$

We can actually cheat here using Bochner's theorem with linearity, and $\Re\varphi(0) = 1$, but using Bochner's theorem can feel like bad practice. Instead let Y be a random variable independent of X with $P(Y = 1) = P(Y = -1) = \frac{1}{2}$, then

$$\varphi_{YX}(t) = E[e^{itYX}] = E[e^{itYX}(1_{Y=1} + 1_{Y=-1})] = E[e^{itX}1_{Y=1}] + E[e^{-itX}1_{Y=-1}]$$

$$\overset{\text{indep}}{=} \varphi(t)E[1_{Y=1}] + \overline{\varphi(t)}E[1_{Y=-1}] = \Re\varphi(t)$$

$\square$

16. (Durrett 3.3.2) (i)

$$\int_{-T}^T e^{-ita}\varphi(t)dt = \int_{-T}^T \int e^{it(X-a)} d\mu dt \overset{\text{Fubini}}{=} \int \int_{-T}^T e^{it(X-a)} dt d\mu = \int \int_{-T}^T (\cos + i\sin)(t(X-a)) dt d\mu$$

$$= \int \int_{-T}^T \cos(t(X-a)) dt d\mu = \int \frac{2\sin(T(X-a))}{X-a} d\mu$$

Fubini here is justified, since the integrand has absolute value ≤ 1, and μ is a probability measure, now applying the limit we want to use DCT, which is justified by $\frac{\sin(T(X-a))}{T(X-a)} \le 1$.

$$\lim_{T\to\infty} \int \frac{2\sin(T(X-a))}{2T(X-a)} d\mu = \int \lim_{T\to\infty} \frac{\sin(T(X-a))}{T(X-a)} d\mu = \int 1_{\{a\}} d\mu = \mu\{a\}$$

(ii) If $x \notin h\mathbb{Z}$, then $P(X = x) = 0$. Now assuming $x \in h\mathbb{Z}$, we find that e^{-itx} is $\frac{2\pi}{h}$ periodic. Moreover

$$\varphi(t + 2\pi/h) = E[\exp(iX(t + 2\pi/h))] = E[1_{h\mathbb{Z}} \exp(iX(t + 2\pi/h))] = \int_{h\mathbb{Z}} \exp(iX(t + 2\pi/h)) d\mu$$

$$= \int_{h\mathbb{Z}} e^{itX} e^{X2\pi/h} d\mu = \int_{h\mathbb{Z}} e^{itX} = \varphi(t)$$

Now, using the previous subpart we find that

$$P(X = x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^T \exp(-itx)\varphi(t)dt$$

Since the limit exists, it suffices to examing a subsequence, using $2\pi/h$ periodicity, we take $T_n = \frac{2\pi n + \pi}{h}$

$$P(X = x) = \lim_{n\to\infty} \frac{h}{2\pi(2n+1)} \int_{-T_n}^{T_n} exp(-itx)\varphi(t)dt = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} exp(-itx)\varphi(t)dt$$

(iii) The equality written in the statement is one of the consequences of the definition of characteristic functions. This result is striaghtforward.

17. (Durrett 3.3.3) $\varphi_{X-Y} = |\varphi|^2$, so by the previous problem, and noting $\{x \mid \mu\{x\} \neq 0\}$ is countable,

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |\varphi(t)|^2 \, dt = P(X - Y = 0) = \int 1_{X-Y=0} d\mu d\mu = \int \int 1_{Y=X} d\mu d\mu = \int \int 1_{Y=x} d\mu d\mu(x)$$

$$= \int \mu\{x\} \, d\mu = \int \mu\{x\} 1_{\mu\{x\}\neq 0} = \sum_{\{x \mid \mu\{x\}\neq 0\}} \mu\{x\} \int 1_{\{x\}} d\mu = \sum_x \mu\{x\}^2$$

$\square$

This exercise has some important consequences, namely if $\varphi(t) \xrightarrow{t\to\infty} 0$, then μ does not have any point masses, moreover the converse is false by the next exercise (3.3.11). There is a partial converse, which says that if μ has a density, the $\varphi(t) \to 0$, to see this if μ has a density f, then $\varphi(t) = \int e^{itx} f(x) dx$, and this is a consequence of the Riemann lebesgue lemma, where we show it for simple functions of intervals, then use that these are dense in $L^1(\mathbb{R})$. On an interval, $\left| \int_a^b e^{itx} \right| = \leq 2/|t|$. $\square$

18. Durrett 3.3.11 This problem provides an example of a random variable, with a continuous density, but not satisfying $\lim_{t\to\infty} \varphi(t) = 0$, the random variable is defined using $X_k \sim \text{Bernoulli}(1/2)$ and $X = 2\sum_1^\infty X_j/3^j$. The distribution function is continuous, since it is the Cantor-Lebesgue function. Now we compute φ, using exercise 3.3.9, the characteristic function for the Bernoulli distribution and $\varphi_{aX}(t) = \varphi_X(at)$ we get

$$\varphi(t) = \prod_1^\infty \varphi_{2X_i/3^j}(t) = \prod_1^\infty \frac{1}{2}\left(e^{it\frac{2}{3^j}} + 1\right)$$

Now when $t = 3^k\pi$, we get

$$\varphi(t) = \prod_1^\infty \frac{1}{2}\left(e^{it\frac{2\pi}{3^{j-k}}} + 1\right) = 1 \cdot \prod_{k+1}^\infty \frac{1}{2}\left(e^{it\frac{2\pi}{3^{j-k}}} + 1\right) = \varphi(\pi)$$

So as long as we can show $|\varphi(\pi)| > 0$, we are done. We first establish

$$\left|e^{i2\pi/3^j} - 1\right| \leq \left|\sin 2\pi/3^j\right| + \left|1 - \cos 2\pi/3^j\right| \leq \frac{2\pi}{3^j} + \frac{2\pi^2}{3^{2j}} \overset{j>2}{\leq} \frac{4\pi}{3^j}$$

Now taking $C = \left|\prod_1^{r-1} \frac{1}{2}\left(e^{it\frac{2}{3^j}} + 1\right)\right|$, where $r > 2$ is chosen so that $2\pi \sum_r^\infty 3^{-j} < \frac{1}{2}$ we find that

$$|\varphi(\pi)| \geq C \prod_r^\infty \frac{1}{2}\left(2 - \frac{4\pi}{3^j}\right) \geq C \prod_r^\infty 1 - \frac{2\pi}{3^j} \geq C - C2\pi \sum_r^\infty 1/3^j \geq C/2 > 0$$

$\square$

19. (Durrett 3.3.4) Take the distribution of a coinflip, with $P(X = 1) = P(X = -1) = \frac{1}{2}$, then

$$\varphi(t) = E[e^{itX}] = \frac{1}{2}(e^{it} - e^{-it}) = \cos t$$

$\cos t \not\to 0$, so it is apparent that $\int_{-\infty}^\infty |\cos t| = \infty$. $\square$

20. (Durrett 3.3.6) Firstly, to solve for the Cauchy distribution, we use that $f(x) = \frac{1}{2\pi} \int e^{-itx} \varphi(t) dt$, in our case we find that $\varphi = e^{-|t|}$, since

$$\frac{2}{\pi} f(x) = \frac{2}{1+x^2} = \frac{1}{1+ix} - \frac{1}{1-ix} = \int_0^\infty e^{-t(1+ix)} dt - \int_0^\infty e^{-t(1-ix)} dt$$

$$= \int_0^\infty e^{-t(1+ix)} dt + \int_{-\infty}^0 e^{-t(ix-1)} dt = \int e^{-|t|} e^{-itx} dt = \int \varphi(t) e^{-itx} dt$$

Then we can use $\varphi_{\frac{\sum_1^n X_k}{n}} = \prod_1^n \varphi_{\frac{X_k}{n}}$, and $\varphi_{\frac{X_k}{n}}(t) = \varphi_{X_k}(t/n)$, so that

$$\varphi_{\frac{\sum_1^n X_k}{n}}(t) = \prod_1^n e^{-|t|/n} = \exp\left(\sum_1^n -|t|/n\right) = e^{-|t|} = \varphi_{X_1}(t)$$

$\square$

21. (Durrett 3.3.7) $X_n \implies X$, hence $\varphi_n \to \varphi$, where $\varphi_n = e^{-\frac{1}{2}\sigma_n^2 t^2}$ and $\varphi = e^{-\frac{1}{2}\sigma^2 t^2}$, noting that $\varphi_n/\varphi \to 1$, we have that $e^{\frac{1}{2}t^2(\sigma^2 - \sigma_n^2)} \to 1$ for all t, now letting $t = \sqrt{2}$, its apparent this only happens when $\sigma_n \to \sigma$. To justify the characteristic functions,

$$E[\exp(itX)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp(itx - \frac{x^2}{2\sigma^2})dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int \cos(tx) e^{-\frac{x^2}{2\sigma^2}} dx$$

Now defining $\varphi(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int \cos(tx) e^{-\frac{x^2}{2\sigma^2}} dx$, we can swap the derivative and integral since our function is absolutely integrable. This gives us

$$\varphi'(t) = \int -x e^{-\frac{x^2}{2\sigma^2}} \sin tx dx$$

Now integrating by parts with $u = \sin tx, du = t \cos tx$, and $dv = -x e^{-\frac{x^2}{2\sigma^2}}$, we get $v = \sigma^2 e^{-\frac{x^2}{2\sigma^2}}$ so that

$$\varphi'(t) = -\sigma^2 t \varphi(t) \implies \varphi(t) = e^{-\frac{\sigma^2 t^2}{2}}$$

$\square$

22. (Durrett 3.3.8) We get $\varphi_n^X \to \varphi^X$ and $\varphi_n^Y \to \varphi^Y$, then $\varphi_n^{X+Y} = \varphi_n^X \varphi_n^Y \to \varphi^X \varphi^Y$.

23. (Durrett 3.3.9) For any n, we get that

$$\varphi_{S_n} = \varphi_{S_{n-1}+X_n} = \varphi_{S_{n-1}} \varphi_{X_n}$$

applying this inductively we get that $\varphi_{S_n} = \prod_1^n \varphi_{X_n}$, since $S_n \to S$, the characteristic functions $\varphi_{S_n} \to \varphi_S$ pointwise, since $S_n \implies S$. Then it follows that for each t, $\varphi_S(t) = \lim_{n\to\infty} \prod_1^n \varphi_{X_n}(t)$, and in particular the limit exists.

24. (Durrett 3.3.12) We first Taylor expand $e^{-t^2/2}$,

$$e^{-t^2/2} = \sum_0^\infty (-1)^n \frac{t^{2n}}{n! 2^n}$$

Now taking $X \sim N(0,1)$, and using the taylor expansion in the last equality, we find that

$$EX^{2n} = \frac{1}{\sqrt{2\pi}} \int x^{2n} e^{-x^2/2} dx = (-1)^n \varphi^{(2n)}(0) = (-1)^n \left(\frac{d}{dt}\right)^{2n}\Bigg|_0 e^{-t^2/2} = \frac{2n!}{2^n n!}$$

$\square$

25. (Durrett 3.3.14) (i) We can't use an L^2 weak law because we only have access to information on the first two terms of the expansion, it will be hard to use a weak law in general due to only having the characteristic functions, so we need to show it directly, first noting

$$\lim_{n\to\infty} \frac{\varphi(t/n) - \varphi(0)}{t/n} = ia \implies n(\varphi(t/n) - 1) = iat$$

we can rearrange to find that

$$\varphi_{S_n/n}(t) = \varphi(t/n)^n = \left(1 + n\frac{\varphi(t/n) - 1}{n}\right)^n$$

So that $\lim_{n\to\infty} \varphi_{S_n/n} = e^{iat} = \varphi_a$ by continuity of log, since e^{iat} is continuous at zero, we get weak convergence $S_n/n \implies a$.

(ii) Immediate consequence of the continuity theorem.

(iii) TODO

26. (Durrett 3.3.16) If $EX = 0$, and $\mathrm{Var}X = 0$, then $EX^2 = 0$, which implies $X \overset{L^2}{=} 0$, so that $X = 0$ a.s.

Now for the main result, we have from taylors theorem $\lim_{h\to 0} \frac{1 - \cos hx}{h^2} = x^2$, this alongside Fatou's lemma gives us

$$E[X^2] \overset{\text{Fatou}}{\leq} 2\liminf_{h\to 0} E\left[\frac{1 - \cos(hX)}{h^2}\right] = 2\liminf_{h\to 0} E\left[\frac{1}{h^2} - \frac{e^{ihx} + e^{-ihx}}{2h^2}\right] = \liminf_{h\to 0}\frac{2 - \varphi(h) - \varphi(-h)}{h^2}$$

$$= -\limsup_{h\to 0}\frac{\varphi(h) + \varphi(-h) - 2}{h^2} = -2\limsup_{h\to 0}\frac{\varphi(h) - 1}{h^2} < \infty$$

Then since $E[X^2] < \infty$, we get that

$$\varphi(t) = 1 + itE[X] - t^2 E[X^2] + o(t^2)$$

so that $E[X] = 0$ for the above limit to make sense. $\qquad\square$

27. (Durrett 3.3.17) One direction is obvious from the continuity theorem, in the other direction $\varphi_n \to \varphi$ where φ is continuous on $(-\delta, \delta)$ so once again by the continuity theorem, $Y_n \implies Y$, where Y has characteristic function φ, we want to show that $\varphi = 1$, so we will be done by the inversion formula. Now since $\lim_{t\downarrow 0}\frac{\varphi(t) - 1}{t^2} = 0 > -\infty$, we get from the proof of the previous exercise (3.3.16) $E[Y^2] = 0$, and hence $Y = 0$ a.s. $\qquad\square$

28. (Durrett 3.3.18) As in the problem statement, it suffices to check that $S_n \to S_\infty$ in probability, since exercise (2.5.10) shows that CIP implies almost sure convergence in this case.

We first would like to show that $(\varphi_n)_n$ is equicontinuous at 0, so let $\epsilon > 0$, writing $\mu_n \sim \varphi_n$, equicontinuity of $(\mu_n)_n$ gives some compact set K such that $\mu_n(K^c) < \epsilon$ for all n, morover since $t \mapsto e^{it}$ is continuous at 0, we can choose δ, so that $|t| < \delta$ implies that $\left|1 - e^{it}\right| < \epsilon$ now letting $M = \sup|K|$, we get for $|t| < \delta/M$

$$|\varphi_n(0) - \varphi_n(t)| \leq \int_K \left|1 - e^{itX}\right| d\mu_n + \int_{K^c} \left|1 - e^{itX}\right| d\mu_n \leq \epsilon + 2\mu_n(K^c) = 3\epsilon$$

establishing equicontinuity.

Now suppose $S_n \implies S_\infty$, then we get $\varphi_n \to \varphi_\infty$, using equicontinuity, we can choose δ such that $t \in N_\delta(0)$ implies $\varphi(t) \geq \frac{1}{2}$, on this neighborhood

$$\lim_{m,n\to\infty} \varphi_{S_n - S_m}(t) = \lim_{m,n\to\infty} \varphi_n(t)/\varphi_m(t) = \varphi_\infty(t)/\varphi_\infty(t) = 1$$

And therefore by exercise (3.3.17) we have $S_n - S_m \implies 0$, to see that it converges to zero in probability it suffices to check it converges to zero in probability along every subsequence $\tilde{S}_k := S_{n_k} - S_{m_k}$. The fact that $\tilde{S}_k \overset{P}{\to} 0$ follows from $\tilde{S}_k \implies 0$ and the fact that zero is constant (c.f. exercise (3.2.12)).

29. (Durrett 3.4.1) The histogram correction in this case is $P(S_{180} \leq 24.5)$, applying the approximation given by the CLT,

$$P(S_{180} \leq 24.5) = P\left(\frac{S_{180} - 30}{5} \leq -11/10\right) = \Phi(-1.1) \approx 13.57$$

$\qquad\square$

30. (Durrett 3.4.4) We may apply the central limit theorem to conclude that $\frac{S_n - n}{\sigma \sqrt{n}} \implies \chi$. Now we consider $B_n = \sqrt{\frac{S_n}{n}} + 1$, then we can apply the Cauchy Schwartz inequality,

$$E\,|X_i| \leq \left(EX_i^2\right)^{\frac{1}{2}} P(\Omega) = \sigma < \infty$$

so that we may apply the strong law of large numbers to $\frac{S_n}{n} \xrightarrow{\text{a.s.}} 1$, it follows that $B_n \to 2$ a.s. and $\frac{1}{B_n} \to \frac{1}{2}$ a.s. which of course implies $\frac{1}{B_n} \implies \frac{1}{2}$, by the converging together lemma for products we get that

$$\frac{S_n - n}{\sigma \sqrt{n}} \frac{1}{B_n} \implies \frac{1}{2}\chi$$

rewriting the left side gives us

$$\frac{(\sqrt{S_n} - \sqrt{n})(\sqrt{S_n} + \sqrt{n})}{\sigma \sqrt{n}} \frac{\sqrt{n}}{(\sqrt{S_n} + \sqrt{n})} \implies \frac{1}{2}\chi$$

Cancelling and multiplying constants gives

$$2(\sqrt{S_n} - \sqrt{n}) \implies \sigma \chi$$

$\square$

31. (Durrett 3.4.5) Let $L_n = \frac{\sigma \sqrt{n}}{\sqrt{\sum_1^n X_k^2}}$ and $R_n = \frac{\sum_1^n X_k}{\sigma \sqrt{n}}$, then $\frac{1}{L_n} \to 1$ a.s. by the strong law, which gives us $L_n \to 1$ a.s. whence $L_n \implies 1$ we also have $R_n \implies \chi$ by the CLT, so by the converging together lemma $L_n R_n \implies \chi$. We are done since

$$T_n = T_n \frac{\sigma \sqrt{n}}{\sigma \sqrt{n}} = L_n R_n$$

$\square$

32. (Durrett 3.4.6) Define $Y_n = \frac{S_{N_n}}{\sigma \sqrt{a_n}}, Z_n = \frac{S_{a_n}}{\sigma \sqrt{a_n}}$, then since $a_n \to \infty$, and $\frac{S_n}{\sigma \sqrt{n}} \implies \chi$ by CLT, we get that $Z_n \implies \chi$. We check now that $Y_n - Z_n \xrightarrow{P} 0$, since this will imply they also converge weakly to zero, then we can use the converging together lemma to get

$$Y_n = Z_n + Y_n - Z_n \implies \chi + 0 = \chi$$

It remains to show convergence in probability, letting $\epsilon > 0$, and choosing $\delta = \frac{\epsilon^3}{2}$, we find that

$$P(|Y_n - Z_n| > \epsilon) = P(|Y_n - Z_n| > \epsilon, \left|1 - \frac{N_n}{a_n}\right| < \delta) + P(|Y_n - Z_n| > \epsilon, \left|1 - \frac{N_n}{a_n}\right| < \delta)$$

$$\leq P(|Y_n - Z_n| > \epsilon, \left|1 - \frac{N_n}{a_n}\right| < \delta) + P(\left|1 - \frac{N_n}{a_n}\right| < \delta)$$

The second term goes to zero by assumption, so it remains to bound the first term.

$$P(|Y_n - Z_n| > \epsilon, \left|1 - \frac{N_n}{a_n}\right| < \delta) \leq P\left(\max_{a_n - \delta \leq m \leq a_n + \delta} \left|\frac{\sum_{\min\{m,a_n\}}^{\max\{m,a_n\}} X_k}{\sigma\sqrt{a_n}}\right| > \epsilon\right)$$

$$= P\left(\max_{a_n - \delta \leq m \leq a_n + \delta} \left|\sum_{\min\{m,a_n\}}^{\max\{m,a_n\}} X_k\right| > \epsilon\sigma\sqrt{a_n}\right)$$

$$\overset{\text{max. ineq.}}{\leq} \frac{1}{\sigma^2 a_n \epsilon^2} \text{Var} \sum_{\lceil a_n(1-\delta)\rceil}^{\lfloor a_n(1+\delta)\rfloor} X_k$$

$$= \frac{1}{\sigma^2 a_n \epsilon^2} |\lfloor a_n(1+\delta)\rfloor - \lceil a_n(1-\delta)\rceil| \sigma^2$$

$$\leq \frac{a_n 2\delta\sigma^2}{\sigma^2 a_n \epsilon^2} = \epsilon$$

So that $|X_n - Y_n| \overset{P}{\longrightarrow} 0$. $\qquad\square$

33. (Durrett 3.4.9) The example shows why the assumption $\sum_1^n E[X_{n,k}^2 1_{|X_{n,k}|>e}] \to 0$ is essential in the LF-CLT.

$$\text{Var}(X_k) = \frac{2k^2}{2} + 2(\frac{1-k^{-2}}{2}) = 2 - k^{-2}$$

It follows that $\text{Var}\left(\frac{S_n}{\sqrt{n}}\right) = \frac{1}{n}\sum_1^n 2 - k^{-2} \geq 2 - \frac{1}{n}\sum_1^\infty k^{-2}$ so by the squeeze theorem the quantity converges to 2, this satisfies condition (i) of the LF-CLT for $X_{n,k} = \frac{X_k}{\sqrt{n}}$, intuitively condition (ii) fails because of the mass at $X_k = \pm k$, however, we don't need to check this explicitly since failure of the LF-CLT confirms our suspicions. Let $Y_k = X_k 1_{|X_k|\leq 1}$ and $Y_{n,k} = \frac{1}{\sqrt{n}}Y_k$. Then $P(X_k \neq Y_k) = \frac{1}{k^2}$, so that Borel-Cantelli tells us $P(X_k \neq Y_k \text{ i.o.}) = 0$, now defining $Z_n = \sum_1^n Y_k$, we find that if x is such that $X_k(x) \neq Y_k(x)$ for only finitely many k, then letting N be an index larger than the last occurence,

$$\frac{|S_n - Z_n|(x)}{\sqrt{n}} \leq \frac{1}{\sqrt{n}} \sum_1^N k^2 \overset{n\to\infty}{\longrightarrow} 0$$

So that $n^{-\frac{1}{2}}(S_n - Z_n) \to 0$ a.s. (weak convergence of course follows from this as well). Then if we can show $Z_n \implies \chi$, we will be done by the converging together lemma. Finally, we check that we can apply the Lindenberg-Feller CLT to $Y_{n,k}$, we first check that $\sum_1^n E[Y_{n,k}^2 1_{|Y_{n,k}|>\epsilon}] \to 0$, but this is straightforward since $|Y_k| \leq 1$, so for $\sqrt{n} > 1/\epsilon$ we get $\sum_1^n E[Y_{n,k}^2 1_{|Y_{n,k}|>\epsilon}] = 0$. Finally we check that $\sum_1^n E[Y_{n,k}^2] \to 1$, which allows us to conclude that $Z_n \implies 1$ by the Lindenberg-Feller CLT.

$$\sum_1^n E[Y_{n,k}^2] = \frac{1}{n}\left(n - \sum_1^n k^{-2}\right) \geq 1 - \frac{1}{n}\sum_1^\infty k^{-2} \to 1$$

This sum is also clearly less than or equal to 1 for all n, so were done by the squeeze theorem. $\qquad\square$

34. (Durrett 3.4.10) Define $X_{n,k} = \frac{1}{\sqrt{\text{Var}S_n}}(X_k - EX_k)$, we simply check the conditions for Lindenberg-Feller CLT.
   (i)

$$\sum_1^n EX_{n,k}^2 = \frac{1}{\text{Var}S_n}\sum_1^n \text{Var}(X_k) = \frac{\text{Var}S_n}{\text{Var}S_n} = 1$$

(ii) Let $\epsilon > 0$, since $\lim_{n\to\infty} \text{Var} S_n = \infty$ we can choose N such that $n \geq N$ implies that $\text{Var} S_n \geq \left(\frac{2M}{\epsilon}\right)^2$, then for $n \geq N$, we have

$$|X_{n,k}| = \frac{1}{S_n^{\frac{1}{2}}} |X_k - E[X_k]| \leq \frac{1}{S_n^{1/2}} 2M \leq \epsilon$$

So that $E[X_{n,k}^2 1_{|X_{n,k}|>\epsilon}] = 0$. $\square$

35. (Durrett 3.4.11) Once again, we check the conditions of the Lindenberg-Feller CLT with $X_{n,k} = X_k/\sqrt{n}$.

(i) $\sum_1^n E[(X_k/\sqrt{n})^2] = \frac{1}{n} n E X_1^2 = 1$

(ii) We can use the Borel-Cantelli lemma by noticing for any $\epsilon > 0$, the following is summable

$$P(|X_n|/\sqrt{n} > \epsilon) \leq \frac{1}{(\epsilon\sqrt{n})^{2+\delta}} E|X_n|^{2+\delta} \leq \frac{C\epsilon^{-2}}{n^{1+\delta/2}}$$

Then $P(|X_n|/\sqrt{n} > \epsilon \text{ i.o.}) = 0$, so for some N, we have $n \geq N$ implies that $|X_n|/\sqrt{n} \leq \epsilon$ almost surely. It follows that for $n \geq N$, we have

$$\sum_1^n E[X_k^2/\sqrt{n} 1_{|X_k|/\sqrt{n}>\epsilon}] = \sum_1^N E[X_k^2/\sqrt{n} 1_{|X_k|/\sqrt{n}>\epsilon}] \leq \frac{1}{n}\sum_1^N E X_k^2 = N/n \to 0$$

$\square$

36. (Durrett 3.4.12 - Lyapunov's Theorem) Once again, we check the conditions of the Lindenberg-Feller CLT with $X_{n,k} = \frac{X_k - E X_k}{\alpha_n}$

(i)

$$\sum_1^n E X_{n,k}^2 = \frac{1}{\alpha_n^2} \sum_1^n \text{Var} X_k = \alpha_n^2/\alpha_n^2 = 1$$

(ii)

$$\sum_1^n E[X_{n,k}^2 1_{|X_{n,k}|>\epsilon}] \leq \sum_1^n E\left[\frac{|X_{n,k}|^\delta}{\epsilon^\delta} X_{n,k}^2 1_{|X_{n,k}|>\epsilon}\right] \leq \frac{1}{\epsilon^\delta}\sum_1^n E\left[|X_{n,k}|^{2+\delta}\right]$$

$$= \frac{1}{\epsilon^\delta} \alpha_n^{-2-\delta} \sum_1^n E[|X_k - E X_k|^{2+\delta}] \to 0$$

The convergence of the above sequence is by assumption. $\square$

37. (Durrett 3.4.13) (a)

$$\sum_1^\infty P(X_k \neq 0) = \sum_1^\infty k^{-\beta} < \infty$$

so by Borel-Cantelli, $P(X_k \neq \text{ i.o.}) = 0$, Considering any x not in the set, let N be the largest index such that $X_N(x) \neq 0$, then for any $n \geq N$, we have

$$(S_\infty - S_n)(x) = \sum_{n+1}^\infty X_k(x) = 0$$

This shows almost sure convergence. $\square$

(b) We can simply check the conditions for the Lindenberg-Feller CLT with $X_{n,k} = X_k/n^{(3-b)/2}$

(i)

$$\sum_1^n E[X_{n,k}^2] = n^{\beta-3} \sum_1^n EX_k^2 = n^{\beta-3} \sum_1^n k^{2-\beta} =: T_n$$

Which is clearly between zero and one for all n. We need to check now that it converges to some $\sigma^2 \in (0,\infty)$, to do so, consider $T_{n+1} - T_n$ (and denote $C = \sum_1^\infty k^{2-\beta} < \infty$), then

$$T_{n+1} - T_n = \frac{1}{n+1} + ((n+1)^{\beta-3} - n^{\beta-3})\sum_1^n k^{2-\beta} \geq \frac{1}{n+1} + ((n+1)^{\beta-3} - n^{\beta-3})C$$

$$\geq \frac{1}{n+1} - Cn^{\beta-3}$$

This is of course positive for sufficiently large n, since $\beta - 3 > 1$, so the sequence T_n is bounded, and eventually increasing hence converges to some $c \in [\inf\{T_n\}, 1]$ by Bolzano Weierstrass. The sequence being eventually increasing, and lying in $(0,1]$ also allows us to conclude $c \geq \inf\{T_n\} > 0$.

(ii) We have

$$|X_{n,k}| \leq \frac{k}{n^{(3-b)/2}} \leq \frac{n}{n^{(3-b)/2}} \leq n^{(\beta-1)/2} \to 0$$

So for sufficiently large n, $|X_{n,k}| \leq \epsilon$ for any $\epsilon > 0$. □

(c) Let $t \in \mathbb{R}$, then consider for $1 \leq k \leq n$, $c_{n,k} := \frac{1-\cos(tk/n)}{k/n}$, then since $\left|\frac{1-\cos x}{x}\right| \leq \min\left\{2/|x|, \frac{|x|}{2}\right\}$, we find that

$$|c_{n,k}| = |t| \left|\frac{1-\cos(tk/n)}{tk/n}\right| \leq \frac{k}{2n} |t^2| \leq \frac{1}{2} |t^2|$$

Noting here that independent of t, we always have $|c_{n,k}|/n \to 0$. Now still for fixed t, consider

$$c := \int_0^1 x^{-1}(1 - \cos xt)dx$$

by a similar inequality to what we just showed, we know that the integrand is bounded on $[0,1]$ thus is Riemann integrable, moreover since the limit of the integrand as $x \to 0$ exists, and it is clearly continuous away from 0, we find that the integrand is continuous, so that the integral is equal to the Riemann sum of any partition, so long as the mesh of the partition goes to zero. From this, we find that

$$\sum_1^n \frac{c_{n,k}}{n} \to c$$

This method of obtaining the inequality unfortunately means we will not be able to use the technique from CLT of writing $|\prod_1^n z_k - \prod_1^n \omega_k| \leq \theta^{n-1} \sum_1^n |z_k - \omega_k|$ since we havent shown $c_{n,k} \to c$, fortunately we can apply a similar proof specific to the real case. Considering n be large enough so that $|c_{n,k}|/n < \frac{1}{2}$

$$\left|-c - \sum_1^n \log(1 - \frac{c_{n,k}}{n})\right| = \left|-c + \sum_1^n \frac{c_{n,k}}{n} + \sum_1^n \eta_{n,k} \frac{c_{n,k}^2}{n^2}\right|$$

Here $|\eta_{n,k}| \leq \sup_{x \in [\frac{1}{2}, \frac{3}{2}]} \frac{1}{2} \frac{1}{(1-x)^2} =: \eta < \infty$ where this estimate on $\eta_{n,k}$ and the above equality follow from Taylor's theorem. It follows that

$$\left|-c - \sum_1^n \log(1 - \frac{c_{n,k}}{n})\right| \leq \left|c - \sum_1^n \frac{c_{n,k}}{n}\right| + \frac{\eta}{n}\left|\sum_1^n c_{n,k}^2/n\right| \leq \left|c - \sum_1^n \frac{c_{n,k}}{n}\right| + \frac{\eta}{n}\sum_1^n \frac{1}{4n}|t|^4$$

$$\leq \left|c - \sum_1^n \frac{c_{n,k}}{n}\right| + \frac{1}{n}\eta |t|^4$$

Which goes to zero as $n \to \infty$. It follows that by continuity of exponential, we get $\prod_1^n 1 - \frac{c_{n,k}}{n} \to e^{-c}$.

Now denote the characteristic function of $\frac{S_n}{n}$ as φ_n, and the characteristic functions of X_k and $\frac{X_k}{n}$ as φ_k and $\varphi_{n,k}$ respectively, then

$$\varphi_n(t) = \prod_1^n \varphi_{n,k}(t) = \prod_1^n \varphi_k(t/n) = \prod_1^n E[\exp(itX_k)] = \prod_1^n 1 - k^{-1} + k^{-1} \cos(tk/n)$$

$$= \prod_1^n 1 - k^{-1}(1 - \cos(tk/n)) = \prod_1^n 1 - n^{-1}\left(\frac{1 - \cos(tk/n)}{k/n}\right) = \prod_1^n 1 - \frac{c_{n,k}}{n}$$

Now we will be done once we show $\varphi(t) = \exp\left(\int_0^1 x^{-1}(1 - \cos xt)dx\right)$ is continuous at $t = 0$ by the continuity lemma. It suffices to check that $\int_0^1 x^{-1}(1 - \cos xt)dx$ is since e is, this is easy to see by the DCT, since at $t = 0$ the integrand is zero, we write

$$\left|x^{-1}(1 - \cos xt)\right| = |t| \left|(tx)^{-1}(1 - \cos xt)\right| \le |tx|^2/2$$

so restricting to $|t| \le 1$ we get the bounded convergence theorem with $\frac{1}{2}$ dominating the function, so that $\lim_{t \to 0} \int_0^1 x^{-1}(1 - \cos xt)dx = \int_0^1 \lim_{t \to 0} x^{-1}(1 - \cos xt)dx = 0$, this gives us that $\varphi(t)$ is continuous at zero, and hence by the continuity theorem $S_n/n \to \aleph$. $\qquad\square$

38. (Durrett 3.6.1) I think the 2 is a mistake, taking any $\delta \in (\frac{1}{2}, 1)$, we could simply take $X = 1$ and $Y = 0$, but then $||\mu - \nu|| \le 1 < 2\delta$ but $P(X \ne Y) > \delta$. I will ignore the 2 and proceed,

$$||\mu - \nu|| = \frac{1}{2}\sum_\omega |\mu(\omega) - \nu(\omega)| = \frac{1}{2}\left(\sum_{\{\omega|\mu(\omega)\ge\nu(\omega)\}} \mu(\omega) - \nu(\omega) + \sum_{\{\omega|\mu(\omega)<\nu(\omega)\}} \nu(\omega) - \mu(\omega)\right)$$

$$\le \frac{1}{2}\left(\sum_{\{\omega|\mu(\omega)\ge\nu(\omega)\}} \mu(\omega) - \mu(\omega)\nu(\omega) + \sum_{\{\omega|\mu(\omega)<\nu(\omega)\}} \nu(\omega) - \mu(\omega)\nu(\omega)\right)$$

$$\le \frac{1}{2}\left(\sum \mu(\omega) - \mu(\omega)\nu(\omega) + \sum \nu(\omega) - \mu(\omega)\nu(\omega)\right)$$

$$= \frac{1}{2}\left(\sum \mu(\omega) - \sum \mu(\omega)\nu(\omega) + \sum \nu(\omega) - \sum \mu(\omega)\nu(\omega)\right)$$

$$= 1 - P(X = Y) = P(X \ne Y)$$

Now in the converse direction, we can define X, Y independent such that they have joint distribution

$$P(X = x, Y = y) = \begin{cases} \min\{\mu(x), \nu(x)\} & x = y \\ \frac{1}{||\mu-\nu||}(\mu(x) - \nu(x))^+(\nu(y) - \mu(y))^+ & x \ne y \end{cases}$$

To check that this indeed satisfies our desideratum we simply compute (note here the condition is symmetric socomputing for X should suffice), firstly suppose $\mu(x) \le \nu(x)$, then

$$P(X = x) = \mu(x) + \sum_{y \ne x} \frac{1}{||\mu - \nu||} 0(\nu(y) - \mu(y))^+ = \mu(x)$$

If $\mu(x) > \nu(x)$ then we can similarly compute

$$P(X = x) = \nu(x) + \sum_{y \ne x} \frac{1}{||\mu - \nu||}(\mu(x) - \nu(x))^+(\nu(y) - \mu(y))^+$$

$$= \nu(x) + \frac{\mu(x) - \nu(x)}{||\mu - \nu||}\sum_y (\nu(y) - \mu(y))^+ = \nu(x) + (\mu(x) - \nu(x))\frac{||\mu - \nu||}{||\mu - \nu||}$$

$$= \mu(x)$$

Now finally we compute

$$P(X \neq Y) = \sum_{x \neq y} \frac{1}{||\mu - \nu||} (\mu(x) - \nu(x))^+ (\nu(y) - \mu(y))^+$$

$$= \frac{1}{||\mu - \nu||} \sum_x \left((\mu(x) - \nu(x))^+ \sum_{y \neq x} (\nu(y) - \mu(y))^+ \right)$$

Now since one of the two terms $(\mu(x) - \nu(x))^+$ and $(\nu(y) - \mu(y))^+$ vanishes when $x = y$, we can reindex the summation to a sum over all y, but then we can recognize the inner sum as $||\mu - \nu||$, continuing the computation we find

$$P(X \neq Y) = \frac{1}{||\mu - \nu||} \sum_x (\mu(x) - \nu(x))^+ ||\mu - \nu|| = \sum_x (\mu(x) - \nu(x))^+ = ||\mu - \nu||$$

So the specified X, Y give the desired conclusion.

39. (Durrett 3.7.1) Let $f(t) := \log P(T > t)$, the memoryless assumption then tells us in this case that $f(t + s) = f(t) + f(s)$, since $P(T > t) = 1 - P(T \leq t)$, and log is continuous we find that f is continuous apart from a countable set. Now we will be done once we prove the following
   If f is continuous at a point and satisfies f(x + y) = f(x) + f(y), then f is linear

Proof. The functional equation gives us for integers $f(z) = zf(1)$, moreover on rationals we get $f(1) = zf(1/z)$, so that $f(1/z) = z^{-1}f(1)$, since this hold on rationals let p be a point of continuity for f, then in some neighborhood of p, we get $f(\lim z_n) = \lim f(z_n)$, since rationals are dense, we get in some neighborhood of p, for any z, $z_n \to z$ and $\lim f(z_n) = \lim z_n f(1) = zf(1)$ whence $f(z) = f(1)z$ in a neighborhood of p, now any z is a rational translate of an element of this neighborhood of p, so for any z we can pick x in a sufficiently small neighborhood of p, and $y \in \mathbb{Q}$ so that $f(z) = f(x + y) = f(x) + f(y) = xf(1) + yf(1) = zf(1)$, so that $f(z)$ is linear. $\square$

40. (Durrett 3.7.2) Let F, G be the respective distribution functions for S, T, then we can use independence

$$P(V > t) = P(S > t, T > t) = P(S > t)P(T > t) = \begin{cases} e^{-\lambda t}e^{-\mu t} = e^{-(\lambda + \mu)t} & t > 0 \\ 1 & t \leq 0 \end{cases}$$

So $V \sim \exp(\lambda + \mu)$. Now we can also compute

$$P(V = S) = P(S \leq T) = \int P(S \leq T(y))dG(y) = \int (1 - e^{-\lambda T(y)})dG(y) = \int_0^\infty (1 - e^{-\lambda y})e^{-\mu y}dy$$

$$= \frac{1}{\mu} - \frac{1}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}$$

$\square$

41. (Durrett 3.7.3) The two identities follow directly from the previous exercise (Durrett 3.7.2) by induction (take $\lambda = \lambda_i$ and $\mu = \sum_{j \neq i} \lambda_j$) since V is exponential μ inductively we are done. The interesting portion of the question is showing independence, if μ, ν denote the probability measures associated to T_i and $\min_{j \neq i} T_j$, then we can use independence and Tolleni's theorem to write

$$P(V > t, I = i) = P(T_i > t, \min_{j \neq i} T_j \geq T_i) = \int \int 1_{s > t} 1_{r \geq s} d\nu(r) d\mu(s) = \int P(\min_{j \neq i} T_j \geq s) 1_{s > t} d\mu(s)$$

$$= \int_t^\infty P(\min_{j \neq i} T_j \geq s) d\mu(s) = \int_t^\infty P(\min_{j \neq i} T_j \geq s) \lambda_i e^{-\lambda_i s} ds$$

Then since $\min_{j \neq i} T_j$ has exponential distribution with parameter $\sum_{j \neq i} \lambda_j$, its distribution function is continuous which gives us $P(\min_{j \neq i} T_j \geq s) = P(\min_{j \neq i} T_j > s) = e^{-s \sum_{j \neq i} \lambda_j}$, this rewrites our integral as

$$\int_t^\infty P(\min_{j \neq i} T_j \geq s) \lambda_i e^{-\lambda_i s} ds = \int_t^\infty \lambda_i e^{-s \sum_1^n \lambda_j} ds = \frac{\lambda_i e^{-t \sum_1^n \lambda_j}}{\sum_1^n \lambda_j} = P(I = i) P(V > t)$$

Since h-intervals and singletons are π-systems generating their respective σ-algebras independence is now a consequence of the $\pi - \lambda$ theorem. $\qquad\square$

42. (Durrett 3.7.4) The start times of calls follow. a Poisson process N with parameter λ, we want to consider the compound Poisson process $\sum_1^{N(t)} Y_j$, where the Y_j represent calls, we will consider this for a fixed time t,

$$Y_j = \begin{cases} 1 & \text{Call ongoing} \\ 0 & \text{Call ended} \end{cases}$$

Then if s is the start time of call, Y_i, we get $P(Y_i = 1) = 1 - G(t - s)$ and $P(Y_i = 0) = G(t - s)$. In this notation the number of active calls is $N_1(t) = \#\{Y_j \mid j \leq N(t) \text{ and } Y_j = 1\}$ (N_0 is defined similarly but for $Y_j = 0$). To find this out, we compute the joint distribution a number of calls have arrived, suppose that $j + k$ calls have arrived by time t, we can denote these calls as $Y_1, \ldots, Y_{j+k}$, these Y_i are iid and moreover for each Y_i, we can assign random variables D_i representing the duration, and Z_i representing the start time. These are also all independent. We want to check that the Z_i are uniformly distributed conditioning on Y_i having arrived,

$$P(Z_i \leq x \mid Y_i \text{ Arrived}) = \frac{P(N(0,x) = 1, N(x,t) = 0)}{P(N(0,t) = 1)} = \frac{e^{-\lambda x} \lambda x e^{-\lambda(t-x)}}{e^{-\lambda t} \lambda t} = x/t$$

Knowing the distribution of Z_i (lets call it F) allows us to compute (since $D_i + Z_i$ is the time call Y_i ends):

$$P(Y_i = 1 \mid Y_i \text{ Arrived}) = P(D_i + Z_i > t) = 1 - P(D_i + Z_i \leq t) = 1 - \int \mathbb{1}_{D_i + Z_i \leq t} dG dF$$

$$= 1 - \int G(t - Z_i) dF = 1 - \int_0^t G(t - s) \frac{ds}{t} = \frac{1}{t} \int_0^t 1 - G(t - s) ds$$

$$= \frac{1}{t} \int_0^t 1 - G(r) dr$$

Now we can copy the proof of Poisson thinning to find that (in order to take care of the t in the exponential I use that since $1 - G(x) + G(x) = 1$ we get $\frac{1}{t} \int_0^t 1 - G(x) + \frac{1}{t} \int_0^t G(x) = t$)

$$P(N_1(t) = k, N_0(t) = j) = P(k \text{ calls ongoing}, j + k \text{ calls arrived})$$
$$= P(N(t) = j + k) P(k \text{ of } Y_i = 1, j \text{ of } Y_i = 0 \mid Y_1, \ldots, Y_{j+k} \text{ arrived})$$
$$= e^{-\lambda t} \frac{(\lambda t)^{j+k}}{(j+k)!} \binom{j+k}{k} P(Y_1 = 1 \mid Y_1 \text{ arrived})^k P(Y_1 = 0 \mid Y_1 \text{ arrived})^j$$
$$= e^{-\lambda t} \frac{(\lambda t)^{j+k}}{(j+k)!} \frac{(j+k)!}{j!k!} \left(\frac{1}{t} \int_0^t 1 - G(r) dr \right)^k \left(1 - \frac{1}{t} \int_0^t 1 - G(r) dr \right)^j$$
$$= e^{-\lambda t} \frac{\lambda^{j+k}}{j!k!} \left(\int_0^t 1 - G(r) dr \right)^k \left(\int_0^t G(r) dr \right)^j$$
$$= e^{-\lambda \int_0^t 1 - G(r) dr} \frac{\left(\lambda \int_0^t 1 - G(r) dr \right)^k}{k!} e^{-\lambda \int_0^t G(r) dr} \frac{\left(\lambda \int_0^t G(r) dr \right)^j}{j!}$$

Then we have that $\sum_0^\infty e^{-\lambda \int_0^t G(r)dr} \frac{\left(\lambda \int_0^t G(r)dr\right)^j}{j!} = 1$, so we can compute using the joint distribution

$$P(N_1(t) = k) = \sum_{j=0}^\infty P(N_1(t) = k, N_0(t) = j)$$

$$= e^{-\lambda \int_0^t 1 - G(r)dr} \frac{\left(\lambda \int_0^t 1 - G(r)dr\right)^k}{k!} \sum_0^\infty e^{-\lambda \int_0^t G(r)dr} \frac{\left(\lambda \int_0^t G(r)dr\right)^j}{j!}$$

$$= e^{-\lambda \int_0^t 1 - G(r)dr} \frac{\left(\lambda \int_0^t 1 - G(r)dr\right)^k}{k!}$$

So $N_1(t)$ is poisson mean $\lambda \int_0^t 1 - G(r)dr$, now for D_i (or any random variable with distribution G), we get

$$\mu = ED_i = \int_0^\infty P(D_i > x)dx = \int_0^\infty 1 - G(x)dx$$

So to show that $\lim_{t\to\infty} N_1(t)$ is poisson distributed with mean $\lambda\mu$, it suffices to show that the distribution functions converge pointwise to Poisson $\lambda\mu$. So let F be the distribution function for Poisson($\lambda\mu$), and F_t be the distribution function for Poisson($\lambda\mu_t$) where $\mu_t = \int_0^t 1 - G(x)dx$ (by definition we have $\mu_t \to \mu$). Then

$$|F(k) - F_t(k)| = \left| \sum_0^k e^{-\lambda\mu} \frac{(\lambda\mu)^j}{j!} - \sum_0^k e^{-\lambda\mu_t} \frac{(\lambda\mu_t)^j}{j!} \right| \le k \max_{0\le j\le k} \frac{\lambda^j}{j!} \left| e^{-\lambda\mu}\mu^j - e^{-\lambda\mu_t}\mu_t^j \right|$$

But since $x \mapsto e^{-\lambda x}x^j$ is continuous, the term $\left| e^{-\lambda\mu}\mu^j - e^{-\lambda\mu_t}\mu_t^j \right| \to 0$ as $t \to \infty$ for each $0 \le j \le k$, since we need only consider finitely many terms we find that $k \max_{0\le j\le k} \frac{\lambda^j}{j!} \left| e^{-\lambda\mu}\mu^j - e^{-\lambda\mu_t}\mu_t^j \right| \to 0$, so that $F_t \to F$ pointwise, which is exactly the statement $N_1(t) \stackrel{t\to\infty}{\Longrightarrow}$ Poisson($\lambda\mu$). $\square$

43. (Durrett 3.7.5) Let X be the number of people attending the basketball game, then $N \sim$ poisson(2190), now let Y_j be the j-th persons birthday, so that $P(Y_j) = k = \frac{1}{365}$ for $k = 1,\ldots,365$, and the Y_j are iid. Now let $N_j = \#\{Y_i \mid i \le N \text{ and } Y_i = j\}$, then by Durrett theorem 3.7.4 we get that the N_j are independent and poisson distributed with rate $2190 P(Y_i = j) = \frac{2190}{365} = 6$. The probability of every birthday having a representative is

$$P(N_1 \ge 1, \ldots, N_{365} \ge 1) = \prod_1^{365} P(N_j \ge 1) = \prod_1^{365} 1 - P(N_j = 0) = (1 - e^{-6})^{365} \approx 0.404$$

$\square$

44. (Durrett 3.7.6) By Durrett theorem 2.1.8 T_k is a sum of exponential random variables with parameter 1, hence has density given by $\frac{x^{k-1}}{(k-1)!}e^{-x}$, it will suffice to show that the density f_n of nV_k^n satisfies $f_n(x) \to \frac{x^{k-1}}{(k-1)!}e^{-x}$ pointwise by Scheffe's theorem (I'll provide a quick proof after the problem). Now we can rewrite $P(nV_k^n \le x)$, by summing over all possibilities where atleast k of the V_j are less than or equal to x. This gives us

$$\sum_{j=k}^n \sum_{A\subset n, \#A=j} P(U_\alpha \le x/n \text{ for } \alpha \in A, U_\beta > x/n \text{ for } \beta \in A^c)$$

Now we can use independence and the fact that they are identically distributed to rewrite as

$$\sum_{j=k}^n \binom{n}{j} (x/n \wedge 1)^j (1 - x/n \wedge 1)^j$$

For sufficiently large n, we have $x/n < 1$, so that we can remove the minimums, since the above is clearly absolutely continuous for every n, each distribution gives rise to a probability density so our application of Scheffe's theorem is fine. This lets us write the distribution as

$$\sum_{j=k}^{n} \binom{n}{j} (x/n)^j (1 - x/n)^j$$

so that the density function is given by

$$\frac{d}{dx} \sum_{j=k}^{n} \binom{n}{j} (x/n)^j (1 - x/n)^j = \sum_{j=k}^{n} \frac{j}{n}\binom{n}{j}(\frac{x}{n})^{j-1}(1 - \frac{x}{n})^{n-j} - \sum_{j=k}^{n} \frac{n-j}{n}\binom{n}{j}(\frac{x}{n})^{j}(1 - \frac{x}{n})^{n-j-1}$$

Now we can rewrite $\frac{j}{n}\binom{n}{j} = \binom{n-1}{j-1}$, and $\frac{n-j}{n}\binom{n}{j} = \binom{n-1}{j}$, using this and reindexing the first summation by replacing j with $j+1$, we get

$$\sum_{j=k-1}^{n-1} \binom{n-1}{j}(\frac{x}{n})^{j}(1 - \frac{x}{n})^{n-j-1} - \sum_{j=k}^{n} \binom{n-1}{j}(\frac{x}{n})^{j}(1 - \frac{x}{n})^{n-j-1}$$

Now all the terms cancel apart from the $j = k-1$ term and the $j = n$ term, $\binom{n-1}{n} = 0$, so the $j = n$ term is zero. This leaves us with

$$f_n(x) = \binom{n-1}{k-1}\left(\frac{x}{n}\right)^{k-1}\left(1 - \frac{x}{n}\right)^{n-k}$$

We can break this up into terms and compute their limits, so long as the limits exist the limit of the product exists and is equal to the product of the limits.

$$\lim_{n\to\infty} \frac{n-k}{n} = \lim_{n\to\infty} 1 - k/n = 1$$

we can also compute

$$\binom{n-1}{k-1}n^{-(k-1)} = \frac{1}{(k-1)!} \prod_{j=1}^{k-1} \frac{n-j}{n}$$

So we can use the squeeze theorem to conclude,

$$\frac{1}{k!} \geq \lim_{n\to\infty} \binom{n-1}{k-1}n^{-(k-1)} \geq \lim_{n\to\infty} \frac{1}{k!} \prod_{1}^{k-1} \frac{n-k}{n} = \frac{1}{k!} \prod_{1}^{k-1} \lim_{n\to\infty} \frac{n-k}{n} = \frac{1}{k!}$$

we also get that

$$\lim_{n\to\infty} \left(1 - \frac{x}{n}\right)^{n-k} = \lim_{n\to\infty} \left(1 - \frac{x}{n}\right)^{n}\left(1 - \frac{x}{n}\right)^{-k} = e^{-x} \lim_{n\to\infty} \left(1 - \frac{x}{n}\right)^{-k} = e^{-x}$$

Finaly we can conclude that

$$\lim_{n\to\infty} f_n(x) = x^{k-1} \lim_{n\to\infty} \binom{n-1}{k-1}n^{-(k-1)} \lim_{n\to\infty}\left(1 - \frac{x}{n}\right)^{n-k} = \frac{x^{k-1}}{(k-1)!}e^{-x}$$

So that f_n converges pointwise to tthe density of T_k as desired. $\qquad\square$

Scheffe's Theorem Suppose probability densities $f_n \to f$, then $\int |f(x) - f_n(x)| \to 0$, in particular this tells us that $F_n(x) \to F(x)$, since

$$|F(x) - F_n(x)| = \left|\int_{-\infty}^{x} f(x) - \int_{-\infty}^{x} f_n(x)\right| \leq \int_{-\infty}^{x} |f - f_n| \leq \int |f - f_n| \to 0$$

Proof.

$$|f(x) - f_n(x)| = f_n(x) - f(x) + 2\max\{f(x) - f_n(x), 0\}$$

so that

$$\int |f(x) - f_n(x)| = \int f_n(x) - \int f(x) + 2\int \max\{f(x) - f_n(x), 0\} = 2\int \max\{f(x) - f_n(x), 0\}$$

since $\int f = \int f_n = 1$, finally $|\max\{f(x) - f_n(x), 0\}| \le |f(x)|$, so we can apply dct. □

45. (Durrett 4.1.3) Let $x \in \mathbb{R}$, then on all of Ω, we get that

$$(E(X|G) + xE(Y|G))^2 = E(X|G)^2 + x^2 E(Y|G)^2 + 2xE(XY|G) \ge 0$$

So this quadratic has discriminant ≤ 0 on all of Ω, writing down the discriminant this gives us

$$(2E(XY|G))^2 - 4E(X|G)^2 E(Y|G)^2 \le 0$$

and dividing by 4 gives us

$$E(XY|G)^2 \le E(X|G)^2 E(Y|G)^2$$

Finally Jensen's inequality for conditional expectation implies that $E(X|G)^2 \le E(X^2|G)$ and $E(Y|G)^2 \le E(Y^2|G)$, and since both are positive we get the following inequality

$$E(XY|G)^2 \le E(X|G)^2 E(Y|G)^2 \le E(X^2|G)E(Y^2|G)$$

□

46. (Durrett 4.1.7) Since $\mathcal{F} = \{\Omega, \emptyset\}$, $E(Z|\mathcal{F})$ must be a constant function for any random variable Z, since any non-constant function is not $\mathcal{F}$ measurable. Because of this,

$$\mathrm{Var}(X|\mathcal{F}) = P(\Omega)\mathrm{Var}(X|\mathcal{F}) = \int \mathrm{Var}(X|\mathcal{F})$$

Now we can just compute using the property of integrating conditional expectation.

$$E(\mathrm{Var}(X|\mathcal{F})) + \mathrm{Var}(E(X|\mathcal{F})) = \int \mathrm{Var}(X|\mathcal{F}) + \int E(X|\mathcal{F})^2 - \left(\int E(X|\mathcal{F})\right)^2$$

$$= \int E(X^2|\mathcal{F}) - E(X|\mathcal{F})^2 + \int E(X|\mathcal{F})^2 - \left(\int X\right)^2$$

$$= \int E(X^2|\mathcal{F}) - (EX)^2$$

$$= \int X^2 - (EX)^2 = EX^2 - (EX)^2 = \mathrm{Var}X$$

□

47. (Durrett 4.1.9) The main step is a computation, note since $X \in \mathcal{G}$, so is X^2 this gives us

$$E((X-Y)^2|\mathcal{G}) = E(X^2|\mathcal{G}) + E(Y^2|\mathcal{G}) - 2E(XY|\mathcal{G}) = X^2 + E(Y^2|\mathcal{G}) - 2XE(Y|\mathcal{G}) = E(Y^2|\mathcal{G}) - X^2$$

Now we can apply the assumption $EY^2 = EX^2$, by taking expectations to get

$$E[(X-Y)^2] = E[E((X-Y)^2|\mathcal{G})] = E[E(Y^2|\mathcal{G})] - EX^2 = EY^2 - EX^2 = 0$$

Assume for contradiction that $X \ne Y$ a.s. then $A = \{(X-Y)^2 > 0\}$ has $P(A) > 0$, since $A = \bigcup_1^\infty A_n$ where $A_n = \{(X-Y)^2 \ge \frac{1}{n}\}$, we get that $P(A_n) \to P(A)$ by continuity from below, hence for some n we have $P(A_n) > 0$, so that

$$0 = E[(X-Y)^2] \ge E[(X-Y)^2 1_{A_n}] \ge \int_{A_n} \frac{1}{n} = P(A_n)/n > 0$$

which is the desired contradiction. □

48. (Durrett 4.1.10) Let Z be a random variable on $L^1(\mathcal{F})$, and let $X = E(Z|\mathcal{G})$, further assume that $X \stackrel{\text{dist}}{=} Z$, this of course implies that $P(X > x) = P(Z > x)$ for all $x \in \mathbb{R}$. Now we can apply layer cake to get that

$$EZ^+ = \int_0^\infty P(Z > x)dx = \int_0^\infty P(X > x)dx = EX^+$$

Then the property of conditional probability further implies that $E[Z1_{X>0}] = E[X1_{X>0}] = EZ^+$, which implies that $EZ^+ = E[Z^+1_{X>0}]$, so that $Z > 0$ implies $X > 0$ a.s. Moreover, there does not exist a set B of positive measure with $Z < 0$ on B, and $B \subset X > 0$, else we would have

$$EZ^+ = E[Z1_{X>0}] = E[Z1_B] + E[Z1_{\{X>0\}\setminus B}] < E[Z1_{\{X>0\}\setminus B}] \leq E[Z^+1_{\{X>0\}\setminus B}] \leq EZ^+$$

which is a contradiction, hence $Z < 0$ implies that $X \leq 0$ a.s. Finally $P(Z = 0) = P(X = 0)$, and $\int_{X=0} X = \int_{X=0} Z$ alongside $Z > 0$ implying $X > 0$ a.s. gives us that $\{X = 0\} = \{Z = 0\}$ a.s.

Now for any $c \in \mathbb{R}$ we can choose $Z = Y - c$, and since $Y \stackrel{\text{dist}}{=} E[Y|\mathcal{G}]$, we have $Y - c \stackrel{\text{dist}}{=} E[Y - c|\mathcal{G}] = E[Y|G] - c$, applying the above gives us that $Y > c$ if and only if $X := E[Y|G]$ is. It follows that for $q \in \mathbb{Q}$ and $n \in \mathbb{Z}_{>0}$ any set of the form $\left\{ q - \frac{1}{2n} < X \leq q + \frac{1}{2n} \text{ and } Y \notin (q - \frac{1}{2n}, q + \frac{1}{2n}] \right\}$ has probability zero, we can use this to get a contradiction. Suppose that $\{X \neq Y\}$ has positive probability, then $\{X \neq Y\} = \bigcup_1^\infty \left\{ |X - Y| > \frac{1}{n} \right\}$, so that for some n, one of these sets must have positive probability. Then

$$\left\{ |X - Y| > \frac{1}{n} \right\} \subset \bigcup_{q \in \mathbb{Q}} \left\{ q - \frac{1}{2n} < X \leq q + \frac{1}{2n} \text{ and } Y \notin (q - \frac{1}{2n}, q + \frac{1}{2n}] \right\}$$

by countable subadditivity one of the sets in the union must have positive probability, but this is a contradiction. So we can conclude indeed $P(X \neq Y) = 0$. $\square$

49. (Durrett 4.2.1) Since $X_1, \ldots, X_n \in \mathcal{G}_n$, so are the σ-algebra they generate, i.e. $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) \subset \mathcal{G}_n$, the properties that $X_n \in \mathcal{F}_n$, and $E|X_n| < \infty$ are immediate by definition of $\mathcal{F}$ and $(X_n, \mathcal{G}_n)$ being a martingale. To check the Martingale property we use that $E(E(X_{n+1}|\mathcal{G}_n)|\mathcal{F}_n) = E(X_{n+1}|\mathcal{F}_n)$, this allows us to write

$$E(X_{n+1}|\mathcal{F}_n) = E(E(X_{n+1}|\mathcal{G}_n)|\mathcal{F}_n) = E(X_n|\mathcal{F}_n) = X_n$$

$\square$

50. (Durrett 4.2.3) First note that for measurable X, Y we can write

$$\max\{X, Y\} = \frac{1}{2}(X + Y) + \frac{1}{2}|X - Y| = \frac{1}{2}(X + Y) + \frac{1}{2}(X - Y)^+ + \frac{1}{2}(Y - X)^+$$

Applying this to $\max\{X_{n+1}, Y_{n+1}\}$ and using linearity of conditional expectation yields

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) = E(X_{n+1}|\mathcal{F}_n) + E(Y_{n+1}|\mathcal{F}_n) + E((X_{n+1} - Y_{n+1})^+|\mathcal{F}_n) + E((Y_{n+1} - X_{n+1})^+|\mathcal{F}_n)$$

Since $X \mapsto X^+$ is convex, we can apply the submartingale property, along with Jensen's to get

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) \geq X_n + Y_n + E((X_{n+1} - Y_{n+1})|\mathcal{F}_n)^+ + E((Y_{n+1} - X_{n+1})|\mathcal{F}_n)^+$$

Now using the fact that $X \mapsto X^+$ is increasing, alongside the submartingale property gives us

$$E((X_{n+1} - Y_{n+1})|\mathcal{F}_n)^+ \geq (X_n - Y_n)^+ \text{ and } E((Y_{n+1} - X_{n+1})|\mathcal{F}_n)^+ \geq (Y_n - X_n)^+$$

We can apply these inequalities to get

$$2E(\max\{X_{n+1}, Y_{n+1}\}|\mathcal{F}_n) \geq X_n + Y_n + (X_n - Y_n)^+ + (Y_n - X_n)^+ = 2\max\{X_n, Y_n\}$$

Dividing out by 2, we get the submartingale property. Now note that since $X_n, Y_n \in \mathcal{F}_n$, and max is borel measurable we have $\max\{X_n, Y_n\} \in \mathcal{F}_n$ and $\max X_n, Y_n \leq |X_n| + |Y_n| \in L^1(\mathcal{F}_n)$, since L^1 is a vectorspace. $\square$

51. (Durrett 4.2.6)
(a) A non-negative Martingale is in particular a nonnegative super-Martingale so by the super-martingale convergence theorem, there is some random variable X_∞, such that $X_n \to X_\infty$. Assume for contradiction that $P(A) > 0$, where $A = \{X_\infty > 0\}$. Then on A, we get that

$$\lim_{n\to\infty} Y_n 1_A = \lim_{n\to\infty} \frac{X_n}{X_{n-1}} 1_A = 1_A \frac{X_\infty}{X_\infty} = 1_A$$

Since $X_\infty \neq 0$ on A. Hence we get that $Y_n \to 1$ on A. Note that by continuity from above, we get that $\lim_{k\to\infty} P(|Y_1 - 1| < \frac{1}{k}) = P(Y_1 = 1) < 1$, the pointwise convergence on A tells us for any $k \in \mathbb{Z}_{>0}$, there exists some $N \in \mathbb{Z}_{>0}$ such that for $n \geq N$, we get $|Y_n - 1| < \frac{1}{k}$ on A. Letting k be arbitrary, and choosing such an N, we get for any $n \geq N$

$$P(A) \leq P\left(|Y_N - 1| < \frac{1}{k}, \ldots, |Y_n - 1| < \frac{1}{k}\right) = P\left(|Y_1 - 1| < \frac{1}{k}\right)^{n-N}$$

So that for any $m \in \mathbb{Z}_{>0}$, we have $P(A) \leq P\left(|Y_1 - 1| < \frac{1}{k}\right)^m$, now since k was arbitrary, and A, m are independent of k, this holds for all k, so taking the limit

$$P(A) \leq \lim_{k\to\infty} P\left(|Y_1 - 1| < \frac{1}{k}\right)^m = P(Y_1 = 1)^m$$

And since this holds for any m, we get that

$$P(A) \leq \lim_{m\to\infty} P(Y_1 = 1)^m = 0$$

which is a contradiction. $\qquad\square$
(b) $\frac{1}{n} \log X_n = \frac{1}{n} \sum_1^n \log Y_j$, by the strong law of large numbers this converges to $E \log Y_1$, so it suffices to check that $E \log Y_1 < 0$. We can rewrite $E \log Y_1 = E[1_{Y_1>1} \log Y_1] + E[1_{Y_1<1} \log Y_1]$, now since $Y_1 \neq 1$ we get $0 < \int_{Y_1>1} Y_1 + \int_{Y_1<1} Y_1 \leq EY_1 = 1$, and moreover since the expectation is 1, if one of the two integrals is nonzero, then so is the other so that

$$0 < \int_{Y_1>1} Y_1, \int_{Y_1<1} Y_1 < 1$$

This means that we can apply the inequality (strict because both values are strictly between $0, 1$)

$$0 = \log(EY_1) \geq \log\left(\int_{Y_1>1} Y_1 + \int_{Y_1<1} Y_1\right) > \log\left(\int_{Y_1>1} Y_1\right) + \log\left(\int_{Y_1<1} Y_1\right)$$

$$\geq \int_{Y_1>1} \log Y_1 + \int_{Y_1<1} \log Y_1 = E[1_{Y_1>1} \log Y_1] + E[1_{Y_1<1} \log Y_1] = E \log Y_1$$

To justify the inequality interchanging the log with the integral, we use Jensen's inequality with exponentiation and the fact log is increasing (i.e. suppose X is a random variable positive on A), then

$$\exp\left(\int_A \log X\right) \leq \int_A \exp \log X = \int_A X$$

so since log is increasing, taking log of both sides preserves the inequality

$$\int_A \log X = \log \exp\left(\int_A \log X\right) \leq \log \int_A X$$

$\qquad\square$

52. (Durrett 4.4.2) Since $N \geq M$, we have $1_{M<n} - 1_{N<n} \geq 0$, so we can take $H_n = 1_{M<n} - 1_{N<n}$ which is positive as above, obviously bounded and also predictable since $\{M < n\} = \{M \leq n-1\} \in \mathcal{F}_{n-1}$ and $\{N < n\} = \{N \leq n-1\} \in \mathcal{F}_{n-1}$, hence $1_{M<n} \in \mathcal{F}_{n-1}$ and $1_{N<n} \in \mathcal{F}_{n-1}$, which implies their difference is also in $\mathcal{F}_{n-1}$. Since H constitutes a positive martingale transform and X_n is a martingale, so is $(H \bullet X)_n$, so that (on the set $N \leq k$ which is a.s.)

$$(H \bullet X)_k = \sum_1^k (1_{M<n} - 1_{N<n})(X_n - X_{n-1}) = \sum_1^k 1_{M<n}(X_n - X_{n-1}) - \sum_1^k 1_{N<n}(X_n - X_{n-1})$$

$$= \sum_1^k (1 - 1_{M\geq n})(X_n - X_{n-1}) + \sum_1^k (1_{N\geq n} - 1)(X_n - X_{n-1})$$

$$= X_k - X_{k\wedge M} + -X_k + X_{k\wedge N} = X_{k\wedge N} - X_{k\wedge M} = X_N - X_M$$

Now since $(H \bullet X)_n$ is a submartingale, we get

$$EX_N - EX_M = E[(H \bullet X)_k] \geq E[(H \bullet X)_0] = 0$$

. $\square$

53. (Durrett 4.4.9) Let $1 \leq m \leq n$, then

$$E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = E[X_m Y_m] + E[X_{m-1}Y_{m-1}] - E[X_m Y_{m-1}] - E[Y_m X_{m-1}]$$

Notice that

$$E[X_m Y_{m-1}] = E[E(X_m Y_{m-1}|\mathcal{F}_{m-1})] = E[Y_{m-1}E(X_m|\mathcal{F}_{m-1})] = E[X_{m-1}Y_{m-1}]$$
$$E[Y_m X_{m-1}] = E[E(Y_m X_{m-1}|\mathcal{F}_{m-1})] = E[X_{m-1}E(Y_m|\mathcal{F}_{m-1})] = E[X_{m-1}Y_{m-1}]$$

so the above expression simplifies to

$$E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = E[X_m Y_m] - E[X_{m-1}Y_{m-1}]$$

of course that means we can telescope the following series

$$\sum_1^n E[(X_m - X_{m-1})(Y_m - Y_{m-1})] = \sum_1^n E[X_m Y_m] - E[X_{m-1}Y_{m-1}] = E[X_n Y_n] - E[X_0 Y_0]$$

$\square$

54. (Durrett 4.6.1) Denote $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$, and $\mathcal{F} = \sigma(\bigcup_1^\infty \mathcal{F}_n)$, then since $\mathcal{F}_n \uparrow \mathcal{F}$, we get $E(\theta|\mathcal{F}_n) \to E(\theta|\mathcal{F})$ a.s. and in L^1 (Durrett theorem 4.6.8). From here the problem will be done if we can show that $\theta \in \mathcal{F}$, since denoting $\phi = E(\theta|\mathcal{F})$, then $A = \{\theta > \phi\} \in \mathcal{F}$ and $B = \{\phi > \theta\} \in \mathcal{F}$, so that we can use the property of conditional expectation to conclude

$$E|\theta - \phi| = \int_A \theta - \phi + \int_B \phi - \theta = \int_A \theta - \theta + \int_B \theta - \theta = 0$$

so indeed $\theta = \phi$ a.s. so that $E(\theta|\mathcal{F}_n) \to \theta$ a.s. and in L^1. It remains to show that $\theta \in \mathcal{F}$, to do so we can use that the pointwise limit of measurable functions is measurable, and realize θ as a pointwise limit of $\mathcal{F}$ measurable functions, since $E|Z_i| < \infty$, and iid, the strong law of large numbers tells us that $\frac{1}{n}\sum_1^n Z_j \to EZ_1$ almost surely, but then this of course implies that

$$S_n := -EZ_1 + \frac{1}{n}\sum_1^n Y_j = -EZ_1 + \frac{1}{n}\sum_1^n Z_j + \theta \to \theta$$

almost surely. Since each S_n is $\mathcal{F}$ measurable, so is its pointwise limit θ. So $\theta \in \mathcal{F}$ and we are done by the argument above. $\square$

1 Poisson Processes

The problems in Durrett relating to Poisson processes are few and highly technical, due to this I feel it will be better preparation to give an overview of the construction of the poisson process, proofs of its properties, as well as the solutions to a few simple application exercises which I pulled from Durrett's other book *Essentials of Stochastic Processes* (EOSP).

Definition 1 (Poisson Process). We say $N(t)$ is a Poisson process rate λ when for $t, s > 0$ and $t_0 < t_1 < \cdots < t_n$

1. $N(t + s) - N(s) \sim \text{Poisson}(\lambda)$

2. $N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$ are independent

Theorem 1 (Existence of Poisson Process). If $X_1, X_2, \ldots$ are iid exponential parameter λ, and $T_n := \sum_1^n X_n$, then $N(t) = \max\{n \mid T_n \leq t\}$ is a Poisson process with parameter λ.

Proof. First recall (or compute) the density of T_n is given by $f_{T_n} = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$

Define $\tilde{T}_k$ depending on t, so that $\tilde{T}_1 = T_{N(t)+1} - t, \tilde{T}_2 = T_{N(t)+2} - t, \ldots$. We first check N_t is Poisson distributed,

$$P(N(t) = n) = P(T_n \leq t < T_{n+1}) = \int_0^t P(X_{n+1} > t - y) f_{T_n}(y) dy$$

$$= \int_0^t e^{-\lambda(t-y)} \lambda e^{-\lambda y} \frac{(\lambda y)^{n-1}}{(n-1)!} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

SO $N(t)$ is Poisson λt. Now since $N(t + s) - N(s)$ is given by a poisson process on the $\tilde{T}_k$, it suffices to check the $\tilde{T}_k$ are independent of $N(t)$ and independent. To check independence from $N(t)$,

$$P(T_{n+1} > s \mid N(t) = n) = P(T_{n+1} > s, T_n \leq t)/P(N(t) = n) = e^{-\lambda(s-t)}$$

which gives independence from $N(t)$. The same computation works for $\tilde{T}_k$ for $k > 1$. To see that the $\tilde{T}_k$ are independent and form a poisson processes it suffices to check that $T'_k = \tilde{T}_k - \tilde{T}_{k-1}$ are independent exponential random variables, but this follows virtually by definition. $\qquad \square \qquad \qquad \square$

Theorem 2 (Sum of Poisson Processes). The sum of independent poisson processes with rates λ and μ is a Poisson process with rate $\lambda + \mu$

Proof. Work with fixed t, then the sum of Poisson random variables is poisson with the sum of the parameters. $\qquad \square$

Theorem 3 (Thinning). If $N(t)$ is a poisson process parameter λ, and $Y_1, Y_2, \ldots$ are iid integer valued random variables, then $N_j(t) = \#\{k \leq N(t) \mid Y_k = j\}$ is a poisson process with rate $\lambda P(Y_1 = j)$.

Proof. The proof is just to compute the joint distribution. $P(N_1(s + t) - N_1(t) = k_1, \ldots, N_r(s + t) - N_r(t) = k_r)$ is the probability of having $\sum_1^r k_j$ arrivals between t and $s + t$ multiplied by the probability of these arrivals being assigned as above. Using the multinomial theorem, the assignment probability is given by $\frac{\sum_1^r k_j!}{\prod_1^r k_j!} \prod_1^r P(Y_1 = j)^{k_j}$ multiplying this with the probability of having $\sum_1^{k_r}$ assignments gives the desired probability. $\qquad \square$

Theorem 4 (Conditioning). Let U_j be iid uniform on $(0, t)$, and V_j^n be the j-th smallest U_j, then conditioning on $N(t) = n$, the following joint distributions are equal

$$(V_1^n, \ldots, V_n^n) \text{ and } (T_1, \ldots, T_n)$$

Proof. We compute the joint density, the joint density for $(X_1, \ldots, X_n)$ is easy to compute by independence as $f(x_1, \ldots, x_n) = \lambda^n \prod_1^n e^{-\lambda x_j}$, then we can simply apply a change of variables $T_k = \sum_1^k X_j$, this is lower triangular with ones on diagonal, hence having jacobian 1, the inverse change of coordinates sends $(t_1, \ldots, t_n) \mapsto (t_1, t_2 - t_1, \ldots, t_n - t_{n-1})$, so applying the change of coordinates gives density $\lambda^n \prod_1^n e^{-\lambda(t_j - t_{j-1})} = \lambda^n e^{-\lambda t_n}$ conditioning on $N(t) = n$ multiplies the numerator by a factor of $e^{-\lambda(t - t_n)}$ and the denominator by $e^{-\lambda t} \frac{(\lambda t)^n}{n!}$ resulting in $\frac{n!}{t^n}$.

To see that this is the joint density of (V_j^n), note that the joint density of (U_k) is $\frac{1}{t^n}$, since each vector of (V_j^n) is giving by summing the rearrangements we get $n!/t^n$ since there are $n!$ permutations. $\qquad \square$

55. (Durrett EOSP 2.1) (a) Exponentially distributed with mean 2 means paramater $\frac{1}{2}$. $P(X > 2) = e^{-1}$
(b)
$$P(X > 5 | X > 3) = e^{-5/2}/e^{-3/2} = e^{-1}$$

$\qquad \square$

56. (Durrett EOSP 2.3) Let X_j be the amount of time it takes person j to catch a fish, then X_j are exponentially distributed with mean $\frac{1}{2}$, so $\lambda = 2$. Now we can find the probability distribution of the maximums, $P(\max\{X_1, X_2, X_3\} \le t)$ by iid is equal to $P(X_1 \le t)^3 = (1 - e^{-\lambda t})^3$, then denote Y as the r.v. representing the maximum.

$$E[Y] = \int_0^\infty P(Y > t)dt = \int_0^\infty 1 - (1 - e^{-\lambda t})^3 = \int_0^\infty 3e^{-\lambda t} - 3e^{-2\lambda t} + e^{-3\lambda t}$$
$$= 3/\lambda + 1/3\lambda - 3/2\lambda = \frac{11}{6\lambda} = \frac{11}{12}$$

$\qquad \square$

57. (Durrett EOSP 2.5) The failures of the batteries defines a Poisson process, and since 2 are in use the rate is $\lambda := 2\frac{1}{\mu} = \frac{1}{50}$. We can let X_k denote the arrival of the k-th failure, so these are exponentially distributed with parameter λ. We want to compute the expectation, we can do it with layer cake.

$$P(X_1 + X_2 > t) = \int_0^\infty P(X_1 > t - y)\lambda e^{-\lambda y}dy = \int_0^t \lambda e^{-\lambda t}dy + \int_t^\infty \lambda e^{-\lambda y}dy = (1 + \lambda t)e^{-\lambda t}$$
$$P(X_1 + X_2 + X_3 > t) = P(X_1 + X_2 > t - X_3) = \int_0^t \lambda(1 + \lambda(t - y))e^{-\lambda(t-y)}e^{-\lambda y}dy + \int_t^\infty \lambda e^{-\lambda y}dy$$
$$\overset{\text{IBP}}{=} \frac{3}{\lambda} = 150$$

$\qquad \square$

58. (Durrett EOSP 2.16) More generally,

$$ET_n = E\left[\sum_1^n X_j\right] = \sum_1^n EX_j = \frac{n}{\lambda}$$
$$E[T_n \mid N(\ell) = k] = \ell + E[T_{n-k}] = \ell + \frac{n - k}{\lambda} \qquad (k < n)$$
$$E[N(n) \mid N(k) = \ell] = \ell + E[N(n - k)] = \ell + \sum_{j=0}^\infty e^{-\lambda(n-k)}\frac{(\lambda(n - k))^j}{j!} = \ell + \lambda(n - k) \qquad (k < n)$$

59. (Durrett EOSP 2.17) (a)

$$P(N(\frac{1}{2}) = 0) = e^{-\frac{1}{2}\lambda} = e^{-\frac{1}{2} \cdot 2} = e^{-1}$$

(b) for $x \ge 0$, $F(x) = e^{-1} + (1 - e^{-1})e^{-2x}$

60. (Durrett EOSP 2.22) Apply change of variables,

$$E[N \circ U] = \int_1^2 E[N(t)]dt = \int_1^2 \lambda t dt = \frac{3}{2}\lambda$$

Now to compute the variance do the same

$$E(N \circ U)^2 - (E[N \circ U])^2 = \int_1^2 E[N(t)]^2 dt - \frac{9}{4}\lambda^2 = \int_1^2 \lambda^2 t^2 dt - \frac{9}{4}\lambda^2 = \lambda^2(\frac{7}{3} - \frac{9}{4}) = \lambda^2/12$$