SOLUTIONS TO HOMEWORK 1

(1) (a) As soon as $\varepsilon < 1$ we have $\mathbb{P}(X_n > \varepsilon) = \mathbb{P}(X_n = 1) = p_n$ so convergence to 0 in probability holds if and only if $p_n \to 0$.

(b) By the first Borel–Cantelli Lemma, if $\sum_{n\geq 1} p_n < \infty$ then

$$\mathbb{P}(X_n = 1 \text{ infinitely often}) = 0$$

so then $X_n \to 0$ almost surely. By the second Borel–Cantelli Lemma, if $\sum_{n\geq 1} p_n = \infty$ then

$$\mathbb{P}(X_n = 1 \text{ infinitely often}) = 1$$

so then X_n does not converge to 0 almost surely.

(2) For $\varepsilon \in (0,1)$ we can write

$$d(X_n,X) = \mathbb{P}(|X_n-X| > 1)+\mathbb{E}[(X_n-X)\mathbb{1}\{|X_n-X| \in (\varepsilon,1]\}]+\mathbb{E}[(X_n-X)\mathbb{1}\{|X_n-X| \leq \varepsilon\}].$$

Assume now that $X_n \to X$ in probability. Then we get

$$d(X_n,X) \leq \mathbb{P}(|X_n - X| > 1) + \mathbb{P}(|X_n - X| > \varepsilon) + \varepsilon \to \varepsilon,$$

so then $d(X_n,X) \to 0$. Conversely, assume $d(X_n,X) \to 0$. Note that for $\varepsilon \in (0,1)$ we have

$$\begin{aligned} d(X_n,X) &\geq \mathbb{P}(|X_n - X| > 1) + \varepsilon\mathbb{P}(|X_n - X| \in (\varepsilon,1])] \\ &\geq \varepsilon\mathbb{P}(|X_n - X| > \varepsilon) \end{aligned}$$

so then $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$ as required.

(3) Fix any constant c. We have

$$\begin{aligned} \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c\}] &= \mathbb{E}[X\mathbb{1}\{Y \leq c\}] - \mathbb{E}[Y\mathbb{1}\{Y \leq c\}] \\ &= \mathbb{E}[\mathbb{E}[X \mid Y]\mathbb{1}\{Y \leq c\}] - \mathbb{E}[Y\mathbb{1}\{Y \leq c\}] = 0, \end{aligned}$$

since $\mathbb{E}[X \mid Y] = Y$. Then

$$\begin{aligned} 0 &= \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X \leq c\}] + \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X > c\}] \\ &\geq \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X \leq c\}], \end{aligned}$$

since the second term is ≥ 0. Similarly, by swapping the roles of X and Y in the indicators,

$$\begin{aligned} 0 &= \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X \leq c\}] + \mathbb{E}[(X-Y)\mathbb{1}\{Y > c, X \leq c\}] \\ &\leq \mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X \leq c\}]. \end{aligned}$$

Taken together, these mean that

$$\mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X \leq c\}] = 0,$$

which in hindsight from the previous equalities means that

$$\mathbb{E}[(X-Y)\mathbb{1}\{Y \leq c, X > c\}] = \mathbb{E}[(X-Y)\mathbb{1}\{Y > c, X \leq c\}] = 0.$$

This is only possible if

$$\mathbb{P}(Y \leq c, X > c) = \mathbb{P}(Y > c, X \leq c) = 0.$$

This is true for any *fixed* c. But if $X \neq Y$ then there is some *rational* q such that either $Y \leq q$ and $X > q$, or vice versa. Thus

$$\mathbb{P}(X \neq Y) \leq \sum_{q \in \mathbb{Q}} \big(\mathbb{P}(Y \leq q, X > q) + \mathbb{P}(Y > q, X \leq q)\big) = 0.$$

(4) (a) The event $\{Y_m \geq r_m\}$ can be written

$$\{Y_m \geq r_m\} = A_m \cap A_{m+1} \cap \cdots \cap A_{m+r_m-1}$$

where $A_m = \{X_m = 0\}$. Thus for any n_0,

$$\{Y_n \geq r_n \text{ i.o.}\} = \bigcap_{n\geq 1} \bigcup_{m \geq n} \{Y_m \geq r_m\} = \bigcap_{n \geq n_0} \bigcup_{m \geq n} \big(A_m \cap A_{m+1} \cap \cdots \cap A_{m+r_m-1}\big)$$

belongs to the tail σ-algebra $\mathcal{H}_\infty$ of the independent events $A_1, A_2, \ldots$ The claim follows from Kolmogorov's 0–1 law.

(b) This is essentially the same as we did in the lectures. We have that

$$\sum_{n\geq 1} \mathbb{P}(Y_n \geq r_n) = \sum_{n \geq 1} \mathbb{P}(Y_n \geq \lceil r_n \rceil) \leq \sum_{n\geq 1} 2^{-r_n} < \infty$$

so the claim follows from the first Borel–Cantelli Lemma BCI.

(c) We may assume that the r_n are all integers, since $\sum_{n \geq 1} 2^{-r_n}/r_n = \infty$ is equivalent to the same condition with r_n replaced by either $\lceil r_n \rceil$ or by $\lfloor r_n \rfloor$. We want to apply BCII but need to work with independent events. For any subsequence n_k we have that

$$\big\{Y_n \text{ i.o.}(n)\big\} \supseteq \big\{Y_{n_k} \text{ i.o.}(k)\big\}$$

where i.o.(k) means 'for infinitely many k's'. We pick our subsequence by $n_1 = 1$ and $n_{k+1} = n_k + r_{n_k}$. Then, since $\{Y_m \geq r\} = A_m \cap A_{m+1} \cap \cdots \cap A_{m+r-1}$, the events $\{Y_{n_k} \geq r_{n_k}\}$ are independent. Then

$$\sum_{k \geq 1} \mathbb{P}(Y_{n_k} \geq r_{n_k}) = \sum_{k \geq 1} 2^{-r_{n_k}} = \sum_{k\geq 1} \frac{n_{k+1} - n_k}{r_{n_k}} 2^{-r_{n_k}} \geq \sum_{k \geq 1} \sum_{n = n_k}^{n_{k+1}-1} \frac{2^{-r_n}}{r_n} = \sum_{n \geq 1} \frac{2^{-r_n}}{r_n} = \infty,$$

where we used that r_n is non-decreasing. The claim follows from BCII.

(d) We have

$$\Big\{ \limsup_{n \to \infty} \tfrac{Y_n}{\log_2 n} \geq 1 \Big\} \supseteq \big\{Y_n \geq \log_2 n \text{ i.o.}\big\}$$

and

$$\sum_{n \geq 1} \frac{2^{-\log_2 n}}{\log_2 n} = \sum_{n\geq 1} \frac{1}{n(\log_2 n)} = \infty.$$

(5) We get

$$\mathbb{E}[X(t)] = 0, \qquad \mathbb{E}[X(t)X(t+h)] = 2\cos(h), \quad \forall\, t,$$
$$\mathbb{E}[X(t)^3] = -2\big(\cos(t)^3 + \sin(t)^3\big).$$

The first two show weak stationarity, the last shows that it is not strictly stationary since the third moment depends on t.

(6) One convenient choice of probability triple would be $\Omega = \mathbb{R}^k$, $\mathcal{F}$ = the Borel product σ-algebra, and

$$\mathbb{P}((a_1, b_1] \times \cdots \times (a_k, b_k]) = \big(F(b_1) - F(a_1)\big)\big(F(b_2) - F(a_2)\big) \cdots \big(F(b_k) - F(a_k)\big),$$

where $F(t) = \mathbb{P}(Z_i \leq t)$. If we use the shift

$$\tau(\omega_1, \omega_2, \dots, \omega_k) = (\omega_2, \omega_3, \dots, \omega_k, \omega_1)$$

and $X(\omega) = \omega_1$ and $X_i(\omega) = X(\tau^{i-1}(\omega))$ then the X_i are distributed as specified.

(a) For any choice of indices $i_1 < i_2 < \cdots < i_n$, the vector $(X_{i_1}, X_{i_2}, \dots, X_{i_n})$ consists of copies of Z which are either independent or identical, and which it is depends only on the differences $i_\ell - i_k$. So it is strictly stationary. (This also follows from the fact that the τ above is measure-preserving.)

We have that, if $n = rk + j$ with $j \in \{0, 1, \dots, k-1\}$ then

$$\frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{r(Z_1 + \cdots + Z_k)}{n} + \frac{Z_1 + \cdots + Z_{j+1}}{n} \to \frac{Z_1 + \cdots + Z_k}{k} \qquad \text{a.s.}$$

Since the limit is not constant, the process cannot be ergodic.

(b) From the ergodic theorem and the above, it must be the case that

$$\mathbb{E}[X_1 \mid \mathcal{I}] = \frac{Z_1 + \cdots + Z_k}{k} = \frac{X_1 + \cdots + X_k}{k} =: W.$$

To check that the right-hand-side agrees with the definition of the left-hand-side, we need to verify (i) $\mathcal{I}$-measurability of W, and (ii) $\mathbb{E}[W \mathbb{1}_A] = \mathbb{E}[X_1 \mathbb{1}_A]$ for all $A \in \mathcal{I}$. For (i), $W(\tau(\omega))) = \frac{X_2 + \cdots + X_k + X_{k+1}}{k} = W$ since $X_{k+1} = X_1$, so W is invariant under shift, meaning it is $\mathcal{I}$-measurable. For (ii),

$$\mathbb{E}[W \mathbb{1}_A] = \frac{1}{k} \sum_{j=1}^{k} \mathbb{E}[X_j \mathbb{1}_A] = \frac{1}{k} \sum_{j=1}^{k} \mathbb{E}[X_j \mathbb{1}_{\tau^{-(j-1)}A}] = \tfrac{1}{k} k \mathbb{E}[X_1 \mathbb{1}_A],$$

by invariance of A and stationarity.

(7) For (a), note first that also $\sum_{k \geq 0} \sigma_k^2 < \infty$ since the $\sigma_k \to 0$. We have $\mathbb{E}[X(t)] = 0$ for all t, and using independence

$$c(h) = \mathrm{Cov}(X(t), X(t+h)) = \mathbb{E}\Big[\sum_{k \geq 0} |C_k|^2 e^{-i\lambda_k (t - (t+h))}\Big] = \sum_{k \geq 0} e^{i\lambda_k h} \sigma_k^2$$

So with $p_k = \sigma_k^2 / \sum_{j \geq 0} \sigma_j^2$ we have

$$\rho(t) = \tfrac{c(t)}{c(0)} = \sum_{k \geq 0} e^{i\lambda_k t} p_k = \mathbb{E}[e^{it\Lambda}]$$

where Λ assigns probability p_k to λ_k.

For (b), we want

$$\mathbb{E}[e^{it\Lambda}] = \exp(\mu(e^{it} - 1))) = \sum_{k \geq 0} e^{-\mu} \tfrac{\mu^k}{k!} e^{ikt}.$$

If we take $Y(t) = X(t)$ as in (a) with $\lambda_k = k$ and $\sigma_k^2 = \mu^k / k!$ then $\sum_{j \geq 0} \sigma_j^s = e^\mu$ and we get what we were after.

(8) (a) The order in which we assign numbers to boxes does not affect the distribution of the outcome. Hence it is exchangeable.

(b) We can imagine that boxes receive labels before numbers are assigned to boxes. Then ξ is the probability of putting a number into a box labelled 1:

$$\xi = \begin{cases} 1, & \text{with probability } p^2, \\ \alpha, & \text{with probability } p(1-p), \\ 1-\alpha, & \text{with probability } p(1-p), \\ 0, & \text{with probability } (1-p)^2. \end{cases}$$

(c) Expanding on the previous idea, ξ will be the sum of the α_k over those k for which the box receives label 1:

$$\xi = \sum_{k\geq 1} \varepsilon_k \alpha_k,$$

where $\varepsilon_k = \mathbb{1}\{\text{box } k \text{ labelled } 1\}$ are independent random variables with distribution $\text{Ber}(p)$.

(9) We assume the following independence: the number of days Y_i between inspections, the events of finding a violation at an inspection, and the fines F_i are all independent. Let S_i be the days when violations are detected; the differences $S_{i+1} - S_i$ are then independent copies of

$$S_1 = \sum_{i=1}^{N_1} Y_i$$

where N_1 is the number of inspections until the first violation is discovered. So N_1 is geometrically distributed with mean $\mathbb{E}[N_1] = \frac{1}{p}$ and using Wald's equation, $\mathbb{E}[S_1] = \mathbb{E}[N_1]\mathbb{E}[Y_1] = 45/p$.

Let $N(t)$ be the number of violations up to day t, so $N(t) = \max\{n : S_n \leq t\}$ is a renewal-process. Let q be the new budget per day, to be determined. Then the accumulated cost up to day t is

$$C(t) = tq + \sum_{i=1}^{N(t)} F_i$$

for which by the renewal-reward theorem the long-term average cost converges to

$$q + \tfrac{p}{45} \cdot 505.$$

We want this to be ≤ 100. (Presumably p will also depend on q.)

(10) (a) We have

$$p_j(t) = q_j(t) + \mathbb{E}\big[\mathbb{P}(X(t-S_1) = j)\mathbb{1}\{S_1 \leq t\}\big] = q_j(t) + \int_0^t p_j(t-s)dF(s)$$

where $F(t) = \mathbb{P}(S_1 \leq t)$, which by Theorem 10.1.11 gives

$$p_j(t) = q_j(t) + \int_0^t q_j(t-s)\,dm(s),$$

where $m(t) = \mathbb{E}[\#\{n \geq 1 : S_n \leq t\}]$.

(b) For $j \in \{0,1,2,3\}$, we have

$$\{X(t) = j,\, S_1 > t\} = \{X_1 + \cdots + X_j \leq t,\, X_1 + \cdots + X_{j+1} > t\}$$

since both events describe that there are exactly j arrivals by time t. Since the arrivals form a Poisson process we then have

$$q_j(t) = \frac{e^{-\lambda t}(\lambda t)^j}{j!}, \qquad j \in \{0,1,2,3\}.$$

Integrating gives (using the superior physics-notation for integrals)

$$\int_0^\infty dt\, e^{-\theta t} q_j(t) = \int_0^\infty dt\, e^{-\theta t} \frac{e^{-\lambda t}(\lambda t)^j}{j!} = \tfrac{1}{\lambda}\big(\tfrac{\lambda}{\lambda+\theta}\big)^{j+1}$$

and

$$\int_0^\infty dt\, e^{-\theta t} \int_0^t dm(s)\, q_j(t-s) = \int_0^\infty dm(s)\, e^{-\theta s} \int_0^\infty dr\, q_j(r) e^{-\theta r}.$$

Now using the density of the Gamma-distribution,

$$m'(s) = \sum_{n\geq 1} \frac{\lambda(\lambda s)^{4n-1} e^{-\lambda s}}{(4n-1)!}$$

we get

$$\begin{aligned}\hat{p}_j(\theta) &= \hat{q}_j(\theta)\Big(1 + \int_0^\infty ds\, m'(s) e^{-\theta s}\Big) = \tfrac{1}{\lambda}\big(\tfrac{\lambda}{\lambda+\theta}\big)^{j+1}\Big(1 + \sum_{n\geq 1}\big(\tfrac{\lambda}{\lambda+\theta}\big)^{4n}\Big) \\ &= \frac{\tfrac{1}{\lambda}\big(\tfrac{\lambda}{\lambda+\theta}\big)^{j+1}}{1 - \big(\tfrac{\lambda}{\lambda+\theta}\big)^4}\end{aligned}$$