Department of Mathematics, IST — Section of Probability and Statistics

Probability Theory

2nd. Exam/ 2nd. Test — 1st. Semester — 2009/10

Duration: 3h/ 1h30m — 2010/02/04 — 5PM, Room Q5.1

- This **exam** has three pages and **eight groups**. The corresponding number of points equals half the ones that are marked, thus, the total number of points of an exam is 20.0.
- If you plan to do the **2nd. test** you should only **solve groups V, VI, VII, VIII**. The total number of points is 20.0.
- Please justify your answers.

Group I — Warm up **2.0 points**

The words "symmetric random walk" refer to this situation. (2.0)

> *The proverbial drunk (PD) is clinging to the lamppost. He decides to start walking. The road runs east and west. In his inebriated state he is as likely to take a step east (forward) as west (backward). In each new position he is again as likely to go forward as backward. Each of his steps are of the same length but of random direction — east or west.*
>
> *http://www.physics.ucla.edu/∼chester/TECH/RandomWalk/3Pane.html*

Admit that each step of PD has length equal to one meter and that he has already taken exactly 100 (a hundred) steps.

Find an approximate value for the probability that PD is within a five meters neighborhood of the lamppost.

- **Process**

  Symmetric random walk (SRW)

- **R.v.**

  Y_n = size of the n^{th} step

  $Y_n \stackrel{i.i.d.}{\sim} Y,\ n \in \mathbb{N}$

  $$P(Y=y) = \begin{cases} p=\frac{1}{2}, & y=1 \text{ step east (forward)} \\ 1-p=\frac{1}{2}, & y=-1 \text{ step west (backward)} \\ 0, & \text{otherwise} \end{cases}$$

  $X_n = \sum_{i=1}^n Y_i$ = position of the PD after n steps ($n \in \mathbb{N}$)

- **Initial condition**

  $X_0 = 0$ (lamppost)

- **Requested probability**

  According to Prop. 0.13 (equation (21)), for large values of n and $a < b$, we have

  $$P(a < X_n \leq b) \simeq \Phi(b/\sqrt{n}) - \Phi(a/\sqrt{n}).$$

  Since X_{100} only takes even values, we get

  $$\begin{aligned} P(-5 \leq X_{100} \leq 5) &= P(-5 < X_{100} \leq 4) \\ &\simeq \Phi(4/\sqrt{100}) - \Phi(-5/\sqrt{100}) \\ &= \Phi(0.4) - [1-\Phi(0.5)] \\ &\stackrel{table}{=} 0.6554 - (1-0.6915) \\ &= 0.3469. \end{aligned}$$

  [Another *reasonable* choice ignoring that X_{100} is an even-valued r.v. (but recalling that is is an integer-valued r.v.): $P(-6 < X_{100} \leq 5)$.]

Group II — Probability spaces and random variables **7.0 points**

1. *Let A and B be two events ($A, B \subset \Omega$).* (2.0)

   Prove that $P(A\Delta B) = P(A^c \Delta B^c)$.[1]

   - **Events**

     $A, B \subset \Omega$

   - **To prove**

     $P(A\Delta B) = P(A^c \Delta B^c)$

   - **Proof**

     On one hand,

     $$\begin{aligned} P(A\Delta B) &\stackrel{Def.\,1.10,\,eq.\,(1.7)}{=} P[(A\backslash B)\cup(B\backslash A)] \\ &= [P(A) - P(A\cap B)] + [P(B) - P(B\cap A)] \\ &= P(A) + P(B) - 2P(A\cap B). \end{aligned}$$

     On the other hand,

     $$\begin{aligned} P(A^c\Delta B^c) &= P[(A^c\backslash B^c)\cup(B^c\backslash A^c)] \\ &= P(A^c) + P(B^c) - 2P(A^c\cap B^c) \\ &= P(A^c) + P(B^c) - 2P[(A\cup B)^c] \\ &= [1-P(A)] + [1-P(B)] - 2\times[1-P(A\cup B)] \\ &= P(A) + P(B) - 2P(A\cap B). \end{aligned}$$

     QED

2. *Suppose $\{A_1, A_2, \cdots\}$ is a sequence of events in $\mathcal{F}$. Prove that* (3.0)

$$P(\liminf A_n) \leq \liminf P(A_n).$$

Briefly discuss the use of the special case of the Fatou's lemma.

- **Sequence of events in $\mathcal{F}$**

  $\{A_1, A_2, \cdots\}$

- **To prove**

  $P(\liminf A_n) \leq \liminf P(A_n)$ (this a part of Prop. 1.71, a special case of the Fatou's lemma!).

- **Proof**

  Note that:

  (i) $\{B_k = \cap_{n=k}^{+\infty} A_n,\ k \in \mathbb{N}\}$ is an increasing sequence of events;

  (ii) $B_k \subseteq A_k$.

  As a result:

$$\begin{aligned} B_k &\uparrow \cup_{k=1}^{+\infty} B_k = \lim_{k\to+\infty} B_k, \text{ by (i) and Prop. 1.27;} \\ P(B_k) &\uparrow P\left(\lim_{k\to+\infty} B_k\right), \text{ by Prop. 1.64;} \\ P(\liminf A_n) &= P\left(\cup_{k=1}^{+\infty} \cap_{n=k}^{+\infty} A_n\right) \\ &= P\left(\cup_{k=1}^{+\infty} B_k\right) \\ &= P\left(\lim_{k\to+\infty} B_k\right) \\ &= \lim_{k\to+\infty} P(B_k) \\ &= \liminf P\left(B_k = \cap_{n=k}^{+\infty} A_n\right) \\ &\overset{(ii)}{\leq} \liminf P(A_k) \\ &= \liminf P(A_n). \end{aligned}$$

  QED

- **Comment**

  The special case of the Fatou's lemma plays a vital role in the proof of continuity of probability functions (Rem. 1.70).

3. *Bob is a high school basketball player. He is a 70% free throw shooter.* (2.0)

   During the season, what is the probability that Bob makes his third free throw on his fifth shot?

- **R.v.**

  X = number of shots until Bob makes his 3rd. free throw

[1] *Recall that Δ represents the symmetric difference between two events.*

- **Distribution**

  $X \sim \text{NegativeBinomial}(r, p)$

- **Parameters**

  $r = 3$

  $p = P(\text{free throw}) = 0.7$

- **P.f.**

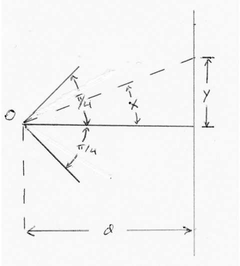
  $P(X = x) = \binom{x-1}{3-1}(1-0.7)^{x-3}\, 0.7^3,\ x = 3, 4, \ldots$

- **Requested probability**

$$\begin{aligned} P(X = 5) &= \binom{5-1}{3-1}(1-0.7)^{5-3}\, 0.7^3 \\ &= \frac{4}{2!\,2!} \times 0.3^2 \times 0.7^3 \\ &= 6 \times 0.09 \times 0.343 \\ &= 0.18522. \end{aligned}$$

Group III — Random variables **7.0 points**

A particle leaves the origin (O) in a free motion, forming an angle X with the horizontal axis, as described in the next figure:



(a) *Let X be a r.v.* (1.5)

Show that $Y = d \times \tan(X)$, $d > 0$ is also a r.v.

- **R.v.**

  X = angle with the horizontal axis

- **Important**

  Let:

  - X be a real r.v.;
  - $(\Omega, \mathcal{F})$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be two measurable spaces.

  Then, by Def. 2.13, $X : \Omega \to \mathbb{R}$ and is such that

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F},\ \forall B \in \mathcal{B}(\mathbb{R}),$$

namely,

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R},$$

according to Prop. 2.16.

- **To prove**
  $Y = g(X) = d \times \tan(X)$ is also a r.v.
- **Proof**
  Firstly, let us remind the reader that a function $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable iff

  $$g^{-1}(B) \in \mathcal{B}(\mathbb{R}), \forall B \in \mathcal{B}(\mathbb{R}),$$

  according to Def. 2.36, with $n = m = 1$. Furthermore, in order that $g : \mathbb{R} \to \mathbb{R}$ be Borel measurable it suffices that $g^{-1}((-\infty, y]) \in \mathcal{B}(\mathbb{R}), \forall y \in \mathbb{R}$, according to Remark 2.47 (with $n = 1$).
  Secondly, note that the range of X is

  $$\mathbb{R}_X = \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$$

  (see (b)), therefore the range of Y is given by

  $$\begin{aligned} \mathbb{R}_Y &= g(\mathbb{R}_X) \\ &= \left[d \times \tan\left(-\frac{\pi}{4}\right), d \times \tan\left(\frac{\pi}{4}\right)\right] \\ &= [-d, d]. \end{aligned}$$

  Thirdly, $g^{-1}((-\infty, y])$ equals:
  - for $y < -d$,
    $$\begin{aligned} \{x \in \mathbb{R} : g(x) \leq y\} &= \emptyset \\ &\in \mathcal{B}(\mathbb{R}); \end{aligned}$$
  - for $-d \leq y \leq d$,
    $$\begin{aligned} \{x \in \mathbb{R} : g(x) \leq y\} &= \{x \in \mathbb{R} : d \times \tan(x) \leq y\} \\ &= \left(-\infty, \arctan\left(\frac{y}{d}\right)\right] \\ &\in \mathcal{B}(\mathbb{R}); \end{aligned}$$
  - for $y > d$,
    $$\begin{aligned} \{x \in \mathbb{R} : g(x) \leq y\} &= \mathbb{R} \\ &\in \mathcal{B}(\mathbb{R}). \end{aligned}$$

  Consequently, we can state that g is indeed a Borel measurable function and therefore Y is a r.v., by Corollary 2.40. QED

(b) *Derive the p.d.f. and the d.f. of Y when $X \sim$ Uniform $\left(-\frac{\pi}{4}, \frac{\pi}{4}\right)$.* (3.0)

- **R.v.**
  X = angle with the horizontal axis
- **Distribution**
  $X \sim \text{Uniform}\left(-\frac{\pi}{4}, \frac{\pi}{4}\right)$

- **P.d.f. of X**
  $$f_X(x) = \begin{cases} \frac{1}{\frac{\pi}{4} - \left(-\frac{\pi}{4}\right)} = \frac{2}{\pi}, & -\frac{\pi}{4} \leq x \leq \frac{\pi}{4} \\ 0, & \text{otherwise} \end{cases}$$
- **Range of X**
  $\mathbb{R}_X = \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$
- **Transformation**
  $Y = g(X) = d \times \tan(X)$
- **Range of Y**
  $\mathbb{R}_Y = g(\mathbb{R}_X) = [-d, d]$
- **Pointwise inverse of g**
  $$\begin{aligned} y &= g(x) \\ &= d \times \tan(x) \\ g^{-1}(y) &= x \\ &= \arctan\left(\frac{y}{d}\right) \end{aligned}$$
- **Derivative of g^{-1}**
  $$\begin{aligned} \frac{d\, g^{-1}(y)}{dy} &= \frac{d \arctan(y/d)}{dy} \\ &= \frac{d(y/d)}{dy} \frac{d \arctan(y/d)}{d(y/d)} \\ &= \frac{1}{d} \frac{1}{1 + (y/d)^2} \\ &= \frac{d}{d^2 + y^2} \end{aligned}$$
- **P.d.f. of Y**
  Since $Y = d \times \tan(X)$ is a strictly increasing function of $x \in \mathbb{R}_X$, we can apply Th. 2.89 and get
  $$\begin{aligned} f_Y(y) &= f_X\left[g^{-1}(y)\right] \times \left|\frac{d\, g^{-1}(y)}{dy}\right| \\ &= f_X\left[\arctan\left(\frac{y}{d}\right)\right] \times \left|\frac{d \arctan(y/d)}{dy}\right| \\ &= \begin{cases} \frac{2}{\pi} \times \left|\frac{d}{d^2+y^2}\right| = \frac{2d}{\pi(d^2+y^2)}, & -d \leq y \leq d \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$
- **D.f. of Y**
  $$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= \int_{-\infty}^{y} f_Y(t)\, dt \\ &= \begin{cases} 0, & y < -d \\ \int_{-d}^{y} \frac{2d}{\pi(d^2+t^2)}\, dt = \cdots = \frac{2}{\pi} \arctan(y/d) + \frac{1}{2}, & -d \leq y \leq d \\ 1, & y > d. \end{cases} \end{aligned}$$
- **Alternatively...**

We could obtain $F_X(x)$, then $F_Y(y)$ in terms of $F_X(x)$ and differentiate $F_Y(y)$ to obtain $f_Y(y)$:

$$\begin{aligned}
F_X(x) &= P(X \le x) \\
&= \int_{-\infty}^{x} f_X(t)\,dt \\
&= \begin{cases} 0, & x < -\frac{\pi}{4} \\ \frac{2}{\pi} \times \left[x - \left(-\frac{\pi}{4}\right)\right] = \frac{2x}{\pi} + \frac{1}{2}, & -\frac{\pi}{4} \le x \le \frac{\pi}{4} \\ 1, & x > \frac{\pi}{4} \end{cases} \\
F_Y(y) &= P(Y \le y) \\
&= P[d \arctan(X) \le y] \\
&= P\left[X \le \tan\left(\frac{y}{d}\right)\right] \\
&= \begin{cases} 0, & y < -d \\ \frac{2 \arctan(y/d)}{\pi} + \frac{1}{2}, & -d \le y \le d \\ 1, & y > d \end{cases} \\
f_Y(y) &= \frac{d\,F_Y(y)}{dy} \\
&= \begin{cases} \frac{d}{dy}\left(\frac{2 \arctan(y/d)}{\pi} + \frac{1}{2}\right) = \cdots = \frac{2d}{\pi(d^2+y^2)}, & -d \le y \le d \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

(c) *Describe a method to generate (pseudo-)random numbers from the distribution of Y.* (2.5)

- **Quantile function**

  Let $u \in (0,1)$ then the quantile function can be derived as follows:

$$\begin{aligned}
F_Y(y) &= u \\
\frac{2 \arctan(y/d)}{\pi} + \frac{1}{2} &= u \\
F_Y^{-1}(u) &= y \\
&= d \times \tan\left[\frac{\pi}{2}\left(u - \frac{1}{2}\right)\right].
\end{aligned}$$

- **Quantile transformation**

  By Prop. 2.140, if

$$U \sim \text{Uniform}(0,1)$$

  then $F_Y^{-1}(U) \stackrel{d}{=} Y$, i.e.,

$$d \times \tan\left[\frac{\pi}{2}\left(U - \frac{1}{2}\right)\right] \stackrel{d}{=} Y.$$

  Hence, if we want to generate (pseudo-)random numbers from the distribution of Y then we have to:

  (1) generate u from a Uniform(0,1) distribution;

  (2) assign $y = d \times \tan\left[\frac{\pi}{2}\left(u - \frac{1}{2}\right)\right]$.

Group IV — Random variables and expectation **4.0 points**

Let X_i be the light field being emitted from a laser at time t_i, $i = 1, 2$.

Laser light is said to be temporally coherent if X_1 and X_2 $(0 < t_1 < t_2)$ are dependent r.v. when $t_2 - t_1$ is not too large.

Admit that the joint p.d.f. of X_1 and X_2 is given by

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\,\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{1-\rho^2}\right),\ x_1, x_2 \in I\!R.$$

(a) *Derive the distribution of $X_1 + X_2$.* (2.0)

- **Random vector**

  $\underline{X} = (X_1, X_2)$

  $X_i =$ light field at time t_i, $i = 1, 2$

- **Distribution**

  Since

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\,\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{1-\rho^2}\right),\ x_1, x_2 \in \mathbb{R},$$

  Exercise 4.184 suggests that (X_1, X_2) has a bivariate normal distribution with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$ and correlation coefficient $\rho \in (-1, 1)$.

- **Transformation**

$$\begin{aligned}
X_1 + X_2 &= \mathbf{C}\underline{X} + \underline{b} \\
&= [1 \quad 1] \times \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\end{aligned}$$

- **Distribution of $X_1 + X_2$**

  According to Th. 4.190,

$$X_1 + X_2 \sim \text{Normal}(\mathbf{C}\,\underline{\mu} + \underline{b}, \mathbf{C}\,\mathbf{\Sigma}\,\mathbf{C}^\top),$$

  where:

$$\begin{aligned}
\mathbf{C}\,\underline{\mu} + \underline{b} &= [1 \quad 1] \times \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
&= 0; \\
\mathbf{C}\,\mathbf{\Sigma}\,\mathbf{C}^\top &= [1 \quad 1] \times \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= [1+\rho \quad \rho+1] \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= 2(1+\rho).
\end{aligned}$$

- **Important**

  Note that the parameters are indeed equal to:

$$\begin{aligned}
E(X_1 + X_2) &= E(X_1) + E(X_2) \\
&= 0 + 0 \\
&= 0;
\end{aligned}$$

$$\begin{aligned}V(X_1+X_2) &= V(X_1)+V(X_2)+2cov(X_1,X_2)\\ &= V(X_1)+V(X_2)+2corr(X_1,X_2)\sqrt{V(X_1)\times V(X_2)}\\ &= 1+1+2\rho\times\sqrt{1\times 1}\\ &= 2(1+\rho).\end{aligned}$$

(b) *Find the d.f. of* $|X_1+X_2|$. (2.0)

- **Another transformation**
  $Y=|X_1+X_2|$
- **Range of** Y
  $\mathbb{R}_Y=\mathbb{R}_0^+$
- **D.f. of** Y

$$\begin{aligned}F_Y(y) &= P(Y\le y)\\ &= \begin{cases}0, & y<0\\ P(-y\le X_1+X_2\le y) & \\ =\Phi\left(\frac{y-E(X_1+X_2)}{\sqrt{V(X_1+X_2)}}\right)-\Phi\left(\frac{-y-E(X_1+X_2)}{\sqrt{V(X_1+X_2)}}\right) & \\ =\Phi\left(\frac{y-0}{\sqrt{2(1+\rho)}}\right)-\Phi\left(\frac{-y-0}{\sqrt{2(1+\rho)}}\right) & \\ =2\,\Phi\left[\frac{y}{\sqrt{2(1+\rho)}}\right]-1, & y\ge 0\end{cases}\end{aligned}$$

Group V — Independence **4.0 points**

Inquiries arrive at a recorded message device according to a Poisson process of rate 15 inquiries per minute.

(a) *Find the probability that in a one-minute period, 3 inquiries arrive during the first 10 seconds and 2 inquiries arrive during the last 15 seconds.* (2.0)

- **Stochastic process**
  $\{N(t),\, t\ge 0\}\sim PP(\lambda=15)$
  $N(t)$ = number of inquiries that arrived at a recorded message device until time t (time in minutes)
  $N(t)\sim \text{Poisson}(\lambda t=15t)$
- **Independent/stationary increments**
  Combining Def. 3.99 with definitions 3.97 and 3.98, it follows that $\{N(t),\, t\ge 0\}$ has:
  – independent increments, thus, $N(a)\perp\!\!\!\perp N(c)-N(b)$, for $0<a<b<c$;
  – stationary increments, therefore $N(b+s)-N(a+s)\stackrel{d}{=}N(b)-N(a)\stackrel{d}{=}N(b-a)\sim \text{Poisson}(\lambda(b-a))$, for $s>0$, $0\le a<b$.

- **Requested probability**

$$\begin{aligned}P[N(1/6)=3,\ N(1)-N(45/60)=2] &\stackrel{indep.\ inc.}{=} P[N(1/6)=3]\\ &\qquad\times P[N(1)-N(45/60)=2]\\ &\stackrel{station.\ inc.}{=} P[N(1/6)=3]\\ &\qquad\times P[N(1-3/4)=2]\\ &\stackrel{N(t)\sim Poisson(\lambda t)}{=} P_{Poisson(15\times\frac{1}{6}=2.5)}(3)\\ &\qquad\times P_{Poisson(15\times\frac{1}{4}=3.75)}(2)\\ &= e^{-2.5}\,\frac{2.5^3}{3!}\times e^{-3.75}\,\frac{3.75^2}{2!}\\ &\simeq 0.035348.\end{aligned}$$

(b) *Admit that 25% of the those inquiries are actually complaints.* (2.0)
If 10 inquiries have arrived to the recorded message device in a one-minute period, what is the probability that at least 3 of those 10 inquiries are complaints?

- **Split process**
  $N_{complaints}(t)$ = number of complaints by time t
  $p=P(\text{inquiry is a complaint})=0.25$
  $\{N_{complaints}(t),t\ge 0\}\sim PP(\lambda p=15\times 0.25=3.75)$, according to Prop. 3.121.
- **Important results**
  $N_{complaints}(t)\sim \text{Poisson}(\lambda p\times t=3.75t)$
  $(N_{complaints}(t)\mid N(t)=n)\sim \text{Binomial}(n,p=0.25)$
- **Requested probability**

$$\begin{aligned}P[N_{complaints}(1)\ge 3\mid N(1)=10] &= 1-P[N_{complaints}(1)\le 2\mid N(1)=10]\\ &= 1-F_{Binomial(10,0.25)}(2)\\ &\stackrel{table}{=} 1-0.5256\\ &= 0.4744.\end{aligned}$$

Group VI — Independence and expectation **7.0 points**

1. *Consider units with independent and exponentially distributed lifetimes with mean equal to one.* (2.5)

   What is the relative change in the survival function when a single-unit system is replaced by a system that operates with two units in parallel? Comment.

   - **R.v.**
     X_i = time to failure of unit i, $i=1,2$
     $X_i\stackrel{i.i.d.}{\sim}X$, $i=1,2$
     $X\sim \text{Exponential}(1)$
     $S(x)=P(X>x)=e^{-x}$, $x\ge 0$

- **New r.v.**
  $X_{(2)} = \max\{X_1, X_2\}$ = lifetime of a parallel system with 2 units
- **Survival function of $X_{(2)}$**
  Capitalizing on Example 3.67, we get

$$\begin{aligned} S_(x) &= P(X_{(2)} > x) \\ &= 1 - [1 - S_X(x)]^2 \\ &= 1 - (1 - e^{-x})^2 \\ &= e^{-x}(2 - e^{-x}), \end{aligned}$$

  for $n = 2$.
- **Requested relative change**

$$\begin{aligned} \frac{S_{X_{(2)}}(x) - S_X(x)}{S_X(x)} \times 100\% &= \frac{e^{-x}(2 - e^{-x}) - e^{-x}}{e^{-x}} \times 100\% \\ &= (1 - e^{-x}) \times 100\% \\ &> 0\%, \; x > 0. \end{aligned}$$

- **Comment**
  Expectedly, there is an improvement in the reliability function when we replace a single-unit system with a parallel system with 2 units.

2. *Suppose that the number of telephone calls made in a day* (X) *is a Poisson r.v. with mean* $\lambda > 0$.

   (a) *Prove that* $E\left(e^{tX}\right) = e^{\lambda(e^t - 1)}$. (1.5)

   - **R.v.**
     X = number of calls in a day
   - **Distribution**
     $X \sim \text{Poisson}(\lambda)$
   - **P.f.**
     $P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$, $x \in \mathbb{N}_0$
   - **Deriving the m.g.f.**
     The m.g.f. (moment generating function) of X is equal to

$$\begin{aligned} E\left(e^{tX}\right) &= \sum_{x=0}^{+\infty} e^{tx} \times e^{-\lambda} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} \sum_{x=0}^{+\infty} \frac{(\lambda e^t)^x}{x!} \\ &= e^{-\lambda} \times e^{\lambda e^t} \\ &= e^{\lambda(e^t - 1)}, \; t \in \mathbb{R}. \end{aligned}$$

QED

   (b) *Find an upper bound for the probability that at least* k *calls* $(k > \lambda)$ *are made in a day using Chernoff's inequality.* (3.0)
   Prove that the minimum upper bound is reached when $t = \ln\left(\frac{k}{\lambda}\right)$.

   - **Upper bound to $P(X \geq k)$**
     Using Chernoff's bound in Prop. 4.99, we can state that

$$\begin{aligned} P(X \geq k) &\leq \frac{E\left(e^{tX}\right)}{e^{tk}} \\ &= \frac{e^{\lambda(e^t - 1)}}{e^{tk}} \\ &= e^{-tk + \lambda(e^t - 1)}. \end{aligned}$$

   - **Minimizing the upper bound**
     Let $g(t) = e^{-tk + \lambda(e^t - 1)}$. Then the smallest value of the upper bound to $P(X \geq k)$ is taken at

$$t^* \; : \; \begin{cases} \left.\frac{d\, g(t)}{dt}\right|_{t=t^*} = 0 \\ \left.\frac{d^2\, g(t)}{dt^2}\right|_{t=t^*} > 0 \end{cases}$$

$$\begin{cases} \left(-k + \lambda e^{t^*}\right) \times g(t^*) = 0 \;\Leftrightarrow\; -k + \lambda e^{t^*} = 0 \;\Leftrightarrow\; t^* = \ln\left(\frac{k}{\lambda}\right) \\ - \end{cases}$$

QED

Group VII — Expectation **7.0 points**

1. *Let* X *be a r.v. such that* $E(X^2) < +\infty$. (3.0)

   Argue and prove that

$$V(X) \geq V(X^-),$$

   by using the fact that $X = X^+ - X^-$, *where* $X^+ = \max\{X, 0\}$ *and* $X^- = -\min\{X, 0\}$ *represent the positive and the negative part of* X *(respectively).*

   - **R.v.**
     $X \in L^2$ (because $E(X^2) < +\infty$)
   - **To prove**
     $V(X) \geq V(X^-)$
   - **Proof**
     Since $X = X^+ - X^-$, where $X^+ = \max\{X, 0\} \geq 0$ and $X^- = -\min\{X, 0\} \geq 0$,

$X^+X^- = 0$ and $E(X^+), E(X^-) \geq 0$, we obtain:

$$\begin{aligned} V(X) &= V(X^+ - X^-) \\ &\overset{(4.112)}{=} V(X^+) + V(X^-) - 2\,cov(X^+, X^-) \\ &\geq V(X^-) - 2\left[E(X^+X^-) - E(X^+)\,E(X^-)\right] \\ &= V(X^-) - 2\left[0 - E(X^+)\,E(X^-)\right] \\ &= V(X^-) + 2\,E(X^+)\,E(X^-) \\ &\geq V(X^-). \end{aligned}$$

QED

- **Argument**
  $X^- = -\min\{X, 0\}$ takes less values than X so it should have a smaller variance than X...

2. *Ten numbers are selected from $\{1, 2, 3, \cdots, 30\}$ uniformly and without replacement.* (4.0)
   Find the expected value of the sum of the selected numbers.[2]

- **R.v.**
  $X_i = i^{th}$ number we pick without replacement from $\{1, \ldots, 30\}$, $i = 1, \ldots, 10$

- **Distribution of X_1**
  Since the selection of the first number is done uniformly,

$$P(X_1 = x_1) = \begin{cases} \frac{1}{30}, & x_1 \in \{1, \ldots, 30\} \\ 0, & \text{otherwise,} \end{cases}$$

  i.e., $X_1 \sim \text{Uniform}(\{1, 2, \ldots, 30\})$.

- **Distribution of X_2**
  Applying the total probability law and capitalising on the fact that the selection is done uniformly and without replacement, we get

$$\begin{aligned} P(X_2 = x_2) &= \sum_{x_1=1}^{30} P(X_2 = x_2 \mid X_1 = x_1) \times P(X_1 = x_1) \\ &= \sum_{x_1=1}^{30} P\left(\text{pick number } x_2 \text{ out of } \{1, \ldots, 30\}\backslash\{x_1\}\right) \times \frac{1}{30} \\ &= \sum_{x_1=1, x_1 \neq x_2}^{30} \frac{1}{29} \times \frac{1}{30} \\ &= 29 \times \frac{1}{29} \times \frac{1}{30} \\ &= \frac{1}{30}, \end{aligned}$$

  that is, $X_2 \sim \text{Uniform}(\{1, 2, \ldots, 30\})$.

[2] *Let X_i be the $i-$th number you pick ($i = 1, \ldots, 10$). Derive the p.f. of X_1, and then the one of X_2 using the total probability law and your knowledge about X_1. Try to generalize this result for X_3, etc.*

- **Distribution of X_3**
  Similarly, applying the multiplication rule:

$$\begin{aligned} P(X_3 = x_3) &= \sum_{x_2=1}^{30} \sum_{x_1=1}^{30} P(X_3 = x_3 \mid X_2 = x_2, X_1 = x_1) \\ &\quad \times P(X_2 = x_2 \mid X_1 = x_1) \times P(X_1 = x_1) \\ &= \sum_{x_2=1,\, x_2 \neq x_3,\, x_1 \neq x_2}^{30} \quad \sum_{x_1=1,\, x_1 \neq x_2}^{30} \frac{1}{28} \times \frac{1}{29} \times \frac{1}{30} \\ &= 28 \times 29 \times \frac{1}{28} \times \frac{1}{29} \times \frac{1}{30} \\ &= \frac{1}{30}, \end{aligned}$$

  i.e., $X_3 \sim \text{Uniform}(\{1, 2, \ldots, 30\})$.

- **Distribution of X_i**
  $X_i \sim \text{Uniform}(\{1, 2, \ldots, 30\})$, $i = 1, \ldots, 10$

- **Expected value of the sum of the 10 selected numbers**
  The linearity of expectation and the fact that the X_i are identically distributed to an Uniform$(\{1, 2, \ldots, 30\})$ r.v. yield:

$$\begin{aligned} E\left(\sum_{i=1}^{10} X_i\right) &= \sum_{i=1}^{10} E(X_i) \\ &= 10 \times E\left[\text{Uniform}(\{1, 2, \ldots, 30\})\right] \\ &= 10 \sum_{x=1}^{30} x \times \frac{1}{30} \\ &= 10 \times \frac{1}{30} \times \frac{30(30+1)}{2} \\ &= 10 \times 15.5 \\ &= 155. \end{aligned}$$

Group VIII — Sequences of random variables **2.0 points**

Let $\{X_1, X_2, \ldots\}$ be a sequence of r.v. such that (2.0)

$$P(\{X_n = x\}) = \begin{cases} 1 - \frac{1}{n^r}, & x = 0 \\ \frac{1}{n^r}, & x = n \\ 0, & \text{otherwise,} \end{cases}$$

where $r \geq 2$.

Prove that: $X_n \stackrel{a.s.}{\to} 0$;[3] *but $X_n \stackrel{q.m.}{\not\to} 0$, for $r = 2$.*

[3] *Use the following result: $\sum_{n=1}^{+\infty} P\left(\{|X_n - X| > \epsilon\}\right) < +\infty, \forall \epsilon > 0 \Rightarrow X_n \stackrel{a.s.}{\to} X$.*

- **Sequence of identically distributed r.v.**

  $\{X_1, X_2, \ldots\}$

- **P.f.**

  $$P(X_n = x) = \begin{cases} 1 - \frac{1}{n^r}, & x = 0 \\ \frac{1}{n^r}, & x = n \\ 0, & \text{otherwise} \end{cases}$$

- **Distribution**

  $X_n \stackrel{d}{=} n \times \text{Bernoulli}\left(\frac{1}{n^r}\right)$, for a fixed $r \geq 2$, and for all $n \in \mathbb{N}$

- **Limiting distribution**

  $X \stackrel{d}{=} 0$

- **Sufficient condition for almost sure convergence**

  In accordance to the table just before Section 5.2,

  $$\sum_{n=1}^{+\infty} P(|X_n - X| > \epsilon) < +\infty, \forall \epsilon > 0 \quad \Rightarrow \quad X_n \stackrel{a.s.}{\to} X.$$

  Since $X \stackrel{d}{=} 0$ and $X_n \geq 0$, this sufficient condition for almost sure convergence can be written as follows:

  $$\sum_{n=1}^{+\infty} P(X_n > \epsilon) < +\infty, \forall \epsilon > 0 \quad \Rightarrow \quad X_n \stackrel{a.s.}{\to} 0.$$

- **Checking if $X_n \stackrel{a.s.}{\to} 0$**

  Note that, for $r \geq 2$, $\sum_{n=1}^{+\infty} \frac{1}{n^r} < +\infty$, for $r \geq 2$. Moreover, X_n takes the following values: 0, with probability $1 - \frac{1}{n^r}$; n, with probability $\frac{1}{n^r}$.

  Now, let $[\epsilon]$ represent the integer part of ϵ. Then, for $\epsilon > 0$,

  $$P(X_n > \epsilon) = \begin{cases} P(X_n = n) = \frac{1}{n^r}, & 0 < \epsilon < n \\ 0, & \epsilon \geq n. \end{cases}$$

  As a result,

  $$\begin{aligned} \sum_{n=1}^{+\infty} P(X_n > \epsilon) &= \sum_{n=1}^{[\epsilon]} 0 + \sum_{n=[\epsilon]+1} \frac{1}{n^r} \\ &< +\infty, \end{aligned}$$

  that is, $X_n \stackrel{a.s.}{\to} 0$. QED

- **Proving that $X_n \stackrel{q.m.}{\not\to} 0$, for $r = 2$**

  If we recall Def. 5.15 then we essentially need to prove that

  $$\lim_{n\to+\infty} E(X_n^2) \neq 0.$$

In fact, for $r = 2$,

$$\begin{aligned} \lim_{n\to+\infty} E(X_n^2) &= \lim_{n\to+\infty} \left[0^2 \times \left(1 - \frac{1}{n^2}\right) + n^2 \times \left(\frac{1}{n^2}\right)\right] \\ &= 1 \\ &\neq 0, \end{aligned}$$

i.e., $X_n \stackrel{q.m.}{\not\to} 0$. QED