Department of Mathematics, IST — Section of Probability and Statistics

Probability Theory

1st. Test — 1st. Semester — 2009/10

Duration: 1h30m — 2009/11/07 — 9AM, Room P12

- Please justify your answers.
- This test has two pages and four groups. The total of points is 20.0.

Group I — Warm up **2.0 points**

The random walk hypothesis is due to French economist Louis Bachelier (1870–1946) and asserts that the random nature of a commodity or stock prices cannot reveal trends and therefore current prices are no guide to future prices. Surprisingly, an investor assumes that his/her daily financial score is governed by a symmetric random walk starting at 0. (2.0)

Obtain the corresponding approximate value for the probability that the fraction of time the financial score is spent positive exceeds 50%.

- **Process**

  Symmetric random walk (SRW)

- **R.v.**

  $Y_i =$ size of the i^{th} step

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

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

  $X_n = \sum_{i=1}^n Y_i =$ position or financial score at time n ($n \in \mathbb{N}$)

- **Initial condition**

  $X_0 = 0$

- **New r.v.**

  $$\begin{aligned} \frac{W_n}{n} &\stackrel{Prop. 0.15}{=} \frac{1}{n}\sum_{i=1}^{n} I_{\mathbb{N}}(X_i + X_{i-1}) \\ &= \text{fraction of time the financial score is spent positive} \end{aligned}$$

- **Limit law (arcsine law)**

  $$\lim_{n\to+\infty} P\left(\frac{W_n}{n} \leq x\right) \stackrel{Prop. 0.15}{=} \frac{2}{\pi}\arcsin\sqrt{x}$$

- **Requested probability (approximate value)**

  $$\begin{aligned} \lim_{n\to+\infty} P\left(\frac{W_n}{n} > 0.5\right) &= 1 - \lim_{n\to+\infty} P\left(\frac{W_n}{n} \leq 0.5\right) \\ &= 1 - \frac{2}{\pi}\arcsin\sqrt{0.5} \\ &= 1 - \frac{2}{\pi}\arcsin\left(\frac{\sqrt{2}}{2}\right) \\ &= 1 - \frac{2}{\pi}\times\frac{\pi}{4} \\ &= \frac{1}{2}. \end{aligned}$$

Group II — Probability spaces **7.0 points**

1. *A non-empty collection of subsets of the sample space Ω, say $\mathcal{M}$, is said to be a monotone class if $\mathcal{M}$ is closed under monotone limits, that is,* (2.0)

   - *if $A_n \uparrow$ and $A_n \in \mathcal{M}$ then $\lim_{n\to+\infty} A_n \in \mathcal{M}$, and*
   - *if $A_n \downarrow$ and $A_n \in \mathcal{M}$ then $\lim_{n\to+\infty} A_n \in \mathcal{M}$.*

   Show that a $\sigma-$ algebra is a monotone class.

   - **To prove**

     $$A_n \uparrow \text{ and } A_n \in \mathcal{M} \Rightarrow \lim_{n\to+\infty} A_n \in \mathcal{M}$$
     $$A_n \downarrow \text{ and } A_n \in \mathcal{M} \Rightarrow \lim_{n\to+\infty} A_n \in \mathcal{M}$$

   - **Proof**

     Let $\mathcal{F}$ be a $\sigma-$algebra on Ω and recall Def. 1.38: $\mathcal{F}$ is a

     – non-empty class of subsets of Ω

     closed under

     (i) countable union

     (ii) countable intersection

     (iii) complementation.

     As a consequence:

     $$\begin{aligned} A_n \uparrow \text{ and } A_n \in \mathcal{F} \Rightarrow \lim_{n\to+\infty} A_n &\stackrel{Prop. 1.27}{=} \cup_{n=1}^{+\infty} A_n \\ &\in \mathcal{F} \quad \text{(because of (i))}; \\ A_n \downarrow \text{ and } A_n \in \mathcal{F} \Rightarrow \lim_{n\to+\infty} A_n &\stackrel{Prop. 1.27}{=} \cap_{n=1}^{+\infty} A_n \\ &\in \mathcal{F} \quad \text{(because of (ii))}. \end{aligned}$$

     That is, the $\sigma-$algebra $\mathcal{F}$ is a monotone class, according to the definition in the test.

2. *Prove the continuity of probability functions*[1] *and briefly discuss the use of this property.* (2.5)

- **Auxiliary result — a special case of the Fatou's lemma** (Prop. 1.17)
  Suppose $A_1, A_2, \dots$ is a sequence of events in $\mathcal{F}$. Then

$$P(\liminf A_n) \le \liminf P(A_n) \le \limsup P(A_n) \le P(\limsup A_n). \quad (1)$$

- **To prove**
  $A_n \to A \Rightarrow P(A_n) \to P(A)$ (continuity of probability functions!)

- **Proof**
  If $A_n \to A$ then we get from Def. 1.22:

$$\liminf A_n = \limsup A_n = A.$$

Therefore

$$P(\liminf A_n) = P(\limsup A_n) = P(A). \quad (2)$$

Taking into account (1) and (2) we conclude that

$$\begin{aligned} P(A) &= P(\liminf A_n) \\ &\le \liminf P(A_n) \\ &\le \limsup P(A_n) \\ &\le P(\limsup A_n) \\ &= P(A), \end{aligned}$$

that is,

$$\liminf P(A_n) = \limsup P(A_n) = P(A)$$

which is equivalent to

$$P(A_n) \to P(A).$$

QED

- **Comment**
  The continuity of probability functions allows us to interchange the limit sign and probabilities: $P(\lim_{n\to+\infty} A_n) = \lim_{n\to+\infty} P(A_n)$.
  This interchange is particularly helpful because it is easier to calculate $\lim_{n\to+\infty} P(A_n)$ than $\lim_{n\to+\infty} A_n$ and, thus, the probability of this limiting event.

3. *The event A is said to be repelled by the event B if $P(A \mid B) < P(A)$, and to be attracted by B if $P(A \mid B) > P(A)$.* (2.5)

   Show that if A is attracted by B then A is repelled by B^c.

[1] *If $A_n \to A$ then $P(A_n) \to P(A)$.*

- **To prove**
  A is attracted by $B \Rightarrow A$ is repelled by B^c, i.e.,

$$P(A \mid B) > P(A) \Rightarrow P(A \mid B^c) < P(A)$$

- **Proof**
  Since the event A is said to be attracted by B if

$$P(A|B) > P(A),$$

we have

$$\begin{aligned} P(A \mid B^c) &= \frac{P(A \cap B^c)}{P(B^c)} \\ &= \frac{P(A \backslash B)}{1 - P(B)} \\ &= \frac{P(A) - P(A \cap B)}{1 - P(B)} \end{aligned}$$

$$\begin{aligned} P(A \mid B^c) &= \frac{P(A) - P(A \mid B)P(B)}{1 - P(B)} \\ &< \frac{P(A) - P(A)P(B)}{1 - P(B)} \\ &= \frac{P(A)[1 - P(B)]}{1 - P(B)} \\ &= P(A), \end{aligned}$$

which means by definition that event A is repelled by B^c. QED

Group III — Random variables and independence **9.5 points**

1. *A particle leaves the origin under the influence of the force of gravity g. Its initial velocity v forms a random angle $\frac{X}{2}$ with the horizontal axis, and it crosses this axis at a point distance $Y = a\,\sin(X)$, where $a = \frac{v^2}{g}$.* (4.0)

   Show that the p.d.f. of Y is given by

$$f_Y(y) = \frac{2}{\pi\sqrt{a^2 - y^2}}, \; 0 \le y \le a,$$

   if we assume that v and g are constant and $X \sim Uniform(0, \pi)$.

- **R.v.**
  $X \sim \text{Uniforme}(0, \pi)$

- **P.d.f. of X**

$$f_X(x) = \begin{cases} \frac{1}{\pi}, & 0 \le x \le \pi \\ 0, & \text{otherwise} \end{cases}$$

- **Range of X**
  $I\!R_X = [0, \pi]$
- **New r.v.**
  $Y = g(X) = a \times \sin(X)$
- **Range of Y**
  $I\!R_Y = g(I\!R_X) = [0, a]$
- **Monotonic restrictions**
  $y = g(x) = a \times \sin(X)$ is function with two monotonic restrictions:
  $$y = \begin{cases} g_1(x) = a \times \sin(x), & x \in [0, \pi/2] \\ g_2(x) = a \times \sin(x), & x \in [\pi/2, \pi]. \end{cases}$$
- **Inverses**
  One deals with $n(y) = 2$ inverses (schematics!):
  $$\begin{aligned} g_1^{-1}(y) &= \arcsin\left(\frac{y}{a}\right) \\ g_2^{-1}(y) &= \pi - \arcsin\left(\frac{y}{a}\right) \end{aligned}$$
- **Derivatives**
  $$\begin{aligned} \frac{dg_1^{-1}(y)}{dy} &= \frac{\frac{1}{a}}{\sqrt{1-\left(\frac{y}{a}\right)^2}} \\ &= \frac{1}{\sqrt{a^2-y^2}} \\ \frac{dg_2^{-1}(y)}{dy} &= \frac{-\frac{1}{a}}{\sqrt{1-\left(\frac{y}{a}\right)^2}} \\ &= -\frac{1}{\sqrt{a^2-y^2}} \end{aligned}$$
- **P.d.f. of Y**
  According to Theorem 2.98, we get:
  $$\begin{aligned} f_Y(y) &= \sum_{k=1}^{n(y)} f_X[g_k^{-1}(y)] \times \left|\frac{dg_k^{-1}(y)}{dy}\right| \\ &= \begin{cases} \frac{1}{\pi} \times \left|\frac{1}{\sqrt{a^2-y^2}}\right| + \frac{1}{\pi} \times \left|-\frac{1}{\sqrt{a^2-y^2}}\right| = \frac{2}{\pi\sqrt{a^2-y^2}}, & y \in I\!R_Y = [0, a] \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

QED

2. *The Pareto distribution, named after the Italian economist Vilfredo Pareto, was originally used to model the wealth of individuals, X.*[2]

We say that $X \sim \text{Pareto}(b, \alpha)$ if
$$f_X(x) = \frac{\alpha\, b^\alpha}{x^{\alpha+1}},\ x \geq b,$$
where $b > 0$ is the minimum possible value of X and $\alpha > 0$ is called the Pareto index.

(a) *Consider n individuals with wealths $X_i \stackrel{i.i.d.}{\sim} X$, $i = 1, \ldots, n$.* (3.0)
Identify the survival function of the minimal wealth of these n individuals and comment on the result.

- **R.v.**
  $X_i \stackrel{i.i.d.}{\sim} \text{Pareto}(b, \alpha),\ b, \alpha > 0$
- **Common d.f.**
  $$\begin{aligned} F_{X_i}(x) &= F_{Pareto(b,\alpha)}(x) \\ &= \int_{-\infty}^{x} f_{Pareto(b,\alpha)}(s)\, ds \\ &= \begin{cases} \int_b^x \frac{\alpha b^\alpha}{s^{\alpha+1}}\, ds = -\, b^\alpha s^{-\alpha}\big|_b^x = 1 - \left(\frac{b}{x}\right)^\alpha, & x \geq b \\ 0, & x < b \end{cases} \end{aligned}$$
- **Transformation**
  $Y = \min_{i=1,\ldots,n} X_i$
- **Survival function of Y**
  $$\begin{aligned} S_Y(x) &= P(Y > x) \\ &= P\left(\min_{i=1,\ldots,n} X_i > x\right) \\ &= P\left(X_i > x,\ i = 1, \ldots, n\right) \\ &\stackrel{X_i\ indep}{=} \prod_{i=1}^n P(X_i > x) \\ &= \prod_{i=1}^n [1 - F_{X_i}(x)] \\ &= \begin{cases} \left(\frac{b}{x}\right)^\alpha, & x \geq b \\ 1, & x < b \end{cases} \\ &= S_{Pareto(b,n\alpha)}(x). \end{aligned}$$
  that is to say $Y \sim \text{Pareto}(b, n\alpha)$.
- **Comment**
  Under certain conditions, the family of Pareto distributions is closed under the algebraic operation of minimum.

(b) *Describe a method to generate (pseudo-)random numbers*[3] *from the $\text{Pareto}(b, \alpha)$ distribution.* (2.5)

[2] *The Pareto distribution seemed to show rather well the way that a larger portion of the wealth of any society is owned by a smaller percentage of the people in that society (http://en.wikipedia.org/wiki/Pareto_distribution).*

[3] *The use of random numbers has expanded beyond random sampling or random assignment of treatments. More common uses are now in simulation studies of physical processes or of analytically intractable mathematical expressions (Gentle, 1998, p. 1).*

- **R.v.**
  $X \sim \text{Pareto}(b, \alpha)$
- **D.f.**
  $F_X(x) = \begin{cases} 1 - \left(\frac{b}{x}\right)^\alpha, & x \geq b \\ 0, & x < b \end{cases}$
- **Quantile function**
  Let $u \in (0,1)$ then the quantile function can be derived as follows:

$$\begin{aligned} F_X(x) &= u \\ 1 - \left(\frac{b}{x}\right)^\alpha &= u \\ \left(\frac{b}{x}\right)^\alpha &= 1-u \\ \frac{b}{x} &= (1-u)^{\frac{1}{\alpha}} \\ F_X^{-1}(u) &= x \\ &= b(1-u)^{-\frac{1}{\alpha}}. \end{aligned}$$

- **Quantile transformation**
  According to Prop. 2.140, if

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

  then $F_X^{-1}(U) \stackrel{d}{=} X$, that is,

  $$b(1-U)^{-\frac{1}{\alpha}} \sim \text{Pareto}(b, \alpha).$$

  Consequently, if we want to generate (pseudo-)random numbers from a $\text{Pareto}(b, \alpha)$ distribution then we have to:
  (1) generate u from a $\text{Uniform}(0,1)$ distribution;
  (2) assign $x = b(1-u)^{-\frac{1}{\alpha}}$.

Group IV — Independence **1.5 points**

Can we state the independence criterion for absolutely continuous random variables[4] *as follows?* (1.5)

> The absolutely continuous random variables $X_1, \ldots, X_n$ are independent iff $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^n f_{X_i}(x_i)$, for all $x_1,\ldots,x_n \in \mathbb{R}$.

Please justify your answer.

- **Comment**

  NO, we cannot state the independence criterion for absolutely continuous r.v. as in the test.

[4] *Let $\underline{X} = (X_1,\ldots,X_n)$ be an absolutely continuous random vector. Then $X_1,\ldots,X_n$ are independent iff $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^n f_{X_i}(x_i)$, for all $x_1,\ldots,x_n \in \mathbb{R}$.*

- **Justification**

  According to Remark 2.59, the converse of Prop. 2.58 is not true, that is, the fact that $X_1,\ldots,X_n$ are absolutely continuous r.v. does not imply that $(X_1,\ldots,X_n)$ is an absolutely continuous random vector and therefore there is no guarantee that there is a joint p.d.f. $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ let alone a p.d.f. which can be factorized.