Department of Mathematics, IST — Probability and Statistics Unit

Probability Theory

1st. TEST 1st. Semester — 2010/11

Duration: 1h30m 2011/01/31 — **3PM**, Room P2

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

Group I — Warm up **3.0 points**

Suppose that in an election, candidate A receives a votes and candidate B receives b votes where $a \in \mathbb{N}$, $b \in \mathbb{N}_0$, and $a > b$.

> *Assuming a random ordering of the votes, what is the probability that A is always ahead of B in the vote count?*

This is an historically famous problem known as the "Ballot Problem", that was solved by Joseph Louis Bertrand in 1887.

(a) *Show that the ballot problem is intimately related to asymmetric random walks.* (1.0)

- **Process**
  $\{X_n,\ n = 0, 1, \ldots, a+b\}$
  $X_0 = 0$ (candidate A is zero votes ahead of B before the counting of the votes starts)
  $X_n =$ difference of votes between the candidates A and B at the count of the n^{th} vote
- **Relating the "Ballot Problem" with a random walk**
  $Y_i =$ size of the i^{th} step
  $Y_i \overset{i.i.d.}{\sim} Y,\ i \in \mathbb{N}$
  $$P(Y = y) = \begin{cases} \frac{a}{a+b}, & y = 1 \text{ (if candidate } A \text{ gets the vote we just counted)} \\ \frac{b}{a+b}, & y = -1 \text{ (if candidate } B \text{ gets the vote we just counted)} \\ 0, & \text{otherwise} \end{cases}$$
  $X_n = \sum_{i=1}^{n} Y_i,\ n \in \mathbb{N}$
- **Conclusion**
  Since $a \in \mathbb{N}$, $b \in \mathbb{N}_0$ and $a > b$, $\{X_n,\ n = 0, 1, \ldots, a+b\}$ is indeed an asymmetrical random walk.

(b) *Let $f(a, b)$ denote the probability that A is always ahead of B in the vote count. Condition on the candidate that receives the last vote to prove that* (2.0)

$$f(a, b) = \frac{a}{a+b} \times f(a-1, b) + \frac{b}{a+b} \times f(a, b-1),$$

with $f(1, 0) = 1$.

Hint: *Use the relationship you established in (a).*

- **Requested identity**

The probability that A is always ahead of B in the vote count is equal to:

(i) $f(a-1, b) = P[X_n > 0,\ n = 1, \ldots, (a-1)+b]$, when candidate A receives $a-1$ votes and candidate B receives b votes;

(ii) $f(a, b-1) = P[X_n > 0,\ n = 1, \ldots, a+(b-1)]$, when candidate A receives a votes and candidate B receives $b-1$ votes.

If an additional and final vote is counted, it is either for candidate A with probability $\frac{a}{a+b}$, or for the candidate B with probability $\frac{b}{a+b}$. So the probability of candidate A being ahead throughout the count to the penultimate vote counted and also after the final vote is:

$$\begin{aligned} f(a, b) &= P(X_n > 0,\ n = 1, \ldots, a+b) \\ &= f(a-1, b) \times \frac{a}{a+b} + f(a, b-1) \times \frac{b}{a+b}. \end{aligned}$$

QED

For more details on the "Ballot Problem", the reader is referred to http://en.wikipedia.org/wiki/Bertrand's_ballot_theorem.[1]

Group II — Probability spaces and random variables **5.5 points**

1. *Consider the sample space Ω, and let $\mathcal{F}$ and $\mathcal{H}$ be two $\sigma -$ algebras on Ω.* (1.5)
   Show that $\mathcal{F} \cap \mathcal{H}$, the collection of subsets of Ω lying both in $\mathcal{F}$ and $\mathcal{H}$, is also a $\sigma -$ algebra.

- **$\sigma -$ algebras on Ω**
  $\mathcal{F}, \mathcal{H}$
- **To prove**
  $\mathcal{F} \cap \mathcal{H}$ is also a $\sigma -$ algebra
- **Proof**
  We ought to mention that, according to Def. 1.38, a minimal set of postulates for a non-empty class of subsets $\mathcal{A}$ of Ω to be a $\sigma -$ algebra on Ω is:
  (i) $\Omega \in \mathcal{A}$;
  (ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;
  (iii) $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \cup_{i=1}^{+\infty} A_i \in \mathcal{A}$.
  Hence, we have to prove that all these 3 postulates are true for $\mathcal{A} = \mathcal{F} \cap \mathcal{H}$.
  Note that $\mathcal{F} \cap \mathcal{H}$ is the collection of subsets of Ω lying both in $\mathcal{F}$ and $\mathcal{H}$. Consequently:
  (i) $\Omega \in \mathcal{F}, \mathcal{H} \Rightarrow \Omega \in \mathcal{F} \cap \mathcal{H}$;
  (ii) $A \in \mathcal{F}, \mathcal{H} \Rightarrow A^c \in \mathcal{F}, \mathcal{H} \Rightarrow A^c \in \mathcal{F} \cap \mathcal{H}$;

[1] The result $f(a, b) = \frac{a-b}{a+b}$ was first published by W.A. Whitworth in 1878, but is named after Joseph Louis François Bertrand who rediscovered it in 1887.

(iii) $A_1, A_2, \dots \in \mathcal{F}, \mathcal{H} \Rightarrow \cup_{i=1}^{+\infty} A_i \in \mathcal{F}, \mathcal{H} \Rightarrow \cup_{i=1}^{+\infty} A_i \in \mathcal{F} \cap \mathcal{H}$.

I.e., $\mathcal{A} = \mathcal{F} \cap \mathcal{H}$ is also a $\sigma-$ algebra on Ω. QED

2. *Let $\mathcal{A}$ be an algebra of subsets of Ω and μ be a* **finitely** *additive probability measure on $\mathcal{A}$. (This requires that $\mu(\Omega) = 1$.) Moreover, consider $\{A_n, n = 1, 2, \dots\}$ a sequence of subsets of Ω such that $A_n \in \mathcal{A}$, $n = 1, 2, \dots$.* (2.0)

   Prove that if $A_n \downarrow \emptyset$ then $\mu(A_n) \downarrow 0$.

   Hint: *Consider $A_0 = \Omega$ and use the Cauchy's convergence criterion.*[2]

- **Setting**

  $(\Omega, \mathcal{A})$ is a measurable space

  μ is finitely additive probability measure on $\mathcal{A}$, $\mu : \mathcal{A} \to [0, 1]$, $\mu(\Omega) = 1$

- **To prove**

  $A_n \downarrow \emptyset \Rightarrow \mu(A_n) \downarrow 0$ (monotone continuity for finitely additive probability measures!)

- **Proof**

  Firstly, μ is finitely additive probability measure on $\mathcal{A}$ essentially means that

  $$\mu\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i),$$

  for any $A_1, \dots, A_n$ pairwise disjoint events in $\mathcal{A}$.

  Secondly, recall that $A_n \downarrow \emptyset$ means that (thinking complements!):

  (i) A_n^c is an increasing sequence of events, i.e., $\emptyset = A_0^c \subseteq A_1^c \subseteq A_2^c \subseteq \dots$

  (ii) $\Omega = \lim A_n^c = \limsup A_n^c = \liminf A_n^c$.

  Thirdly, Prop. 1.54, equation (1.41), suggests this probability measure is monotone (prove it!), thus, (i) implies that

  $$\mu(A_n^c) \uparrow .$$

  Moreover, as a result of (i) and (ii)

  $$\Omega \overset{Prop. 1.27}{=} \cup_{i=1}^{+\infty} A_i^c$$

  $$\begin{aligned} \mu(\Omega) &= 1 \\ &= \mu\left(\cup_{i=1}^{+\infty} A_i^c\right). \end{aligned}$$

  Now, applying the "disjointification" technique, the finite additivity of μ and conveniently recognising a telescopic sum with $\mu(A_0^c) = \mu(\emptyset) = 0$ yields:

  $$\begin{aligned} \mu\left(\cup_{i=1}^{+\infty} A_i^c\right) &= \mu\left(\left[\cup_{i=1}^{n} (A_i^c \backslash A_{i-1}^c)\right] \cup \left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right]\right) \\ &= \mu\left[\cup_{i=1}^{n} (A_i^c \backslash A_{i-1}^c)\right] + \mu\left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right] \\ &= \sum_{i=1}^{n} [\mu(A_i^c) - \mu(A_{i-1}^c)] + \mu\left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right] \quad \text{(because } A_{i-1}^c \subseteq A_i^c\text{)} \\ &= \mu(A_n^c) + \mu\left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right]. \end{aligned}$$

[2] *The series $\sum_{i=0}^{+\infty} a_i$ is convergent iff, for any $m \in \mathbb{N}_0$, $\sum_{i=n}^{n+m} a_i \to 0$ when $n \to +\infty$.*

  Finally, using the fact that

  $$\lim_{n\to+\infty} \left\{\mu(A_n^c) + \mu\left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right]\right\} = 1$$

  and the Cauchy's convergence criterion, we can state that

  $$\lim_{n\to+\infty} \mu\left[\cup_{i=n+1}^{+\infty} (A_i^c \backslash A_{i-1}^c)\right] = 0.$$

  and also that

  $$\begin{aligned} \lim_{n\to+\infty} \mu(A_n^c) &= \mu\left(\lim_{n\to+\infty} A_n^c\right) \\ \lim_{n\to+\infty} \mu(A_n) &= \mu\left(\lim_{n\to+\infty} A_n\right) \end{aligned}$$

  Thus, $\mu(A_n) \downarrow$ and $\lim_{n\to+\infty} \mu(A_n) = \mu(\emptyset)$, that is,

  $$\mu(A_n) \downarrow \mu(\emptyset) = 0.$$

  QED

3. *Three brothers and their partners decide to have children until each couple has two female children.* (2.0)

   Determine the probability that the total number of children born to the three couples does not exceed 8, by admitting that:

   - *$p = P($birth of a female$) = 0.5$ at any time and for every couple;*
   - *the couples procreate independently;*
   - *the gender of the next child is independent of the gender(s) of the previous child(ren);*
   - *there is no possibility of twins, triplets, etc.*

- **R.v.**

  $X_i =$ number of children until couple i has two two female children, $i = 1, 2, 3$

- **Distribution**

  $X_i \overset{i.i.d.}{\sim} \text{NegativeBinomial}(r, p)$, $i = 1, 2, 3$

- **Parameters**

  $r = 2$

  $p = P(\text{birth of a female}) = 0.5$

- **New r.v.**

  $Y = \sum_{i=1}^{3} X_i =$ total number of children born to the 3 couples until all have 2 female children

- **Distribution of Y**

  According to Prop. 3.55,

  $$Y \sim \text{NegativeBinomial}(r^*, p^*),$$

where

$$\begin{aligned} r^* &= r + r + r \\ &= 6 \\ p^* &= p \\ &= 0.5. \end{aligned}$$

- **Requested probability**

$$\begin{aligned} P(Y \leq 8) &= F_{NegativeBinomial(6,0.5)}(8) \\ &\overset{form}{=} 1 - F_{Binomial(8,0.5)}(6-1) \\ &\overset{table}{=} 1 - 0.8555 \\ &= 0.1445. \end{aligned}$$

Group III — Random variables **5.0 points**

Let X be the mass/weight (in g) of an industrial article and admit that $X \sim Exponential(\lambda)$, where λ is such that $E(X) = 75$.

The only available scale automatically reduces to $100g$ any mass/weight larger than $100g$. Let Y be the mass/weight shown by this scale.

(a) *Define Y in terms of X and show that Y is a Borel measurable function, therefore also a r.v.* (1.5)

- **R.v.**
  $X =$ mass/weight (in g) of an industrial article
- **Distribution**
  $X \sim$ Exponential(λ)
- **Parameter**

$$\lambda \; : \; E(X) = 75$$
$$\lambda = 75^{-1}$$

- **Transform**

$$\begin{aligned} Y &= \text{mass/weight shown by the scale} \\ &= \begin{cases} X, & \text{if } 0 \leq X \leq 100 \\ 100, & \text{if } X > 100 \end{cases} \\ &= \min\{X, 100\} \\ &= g(X) \end{aligned}$$

- **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

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

  namely, $X^{-1}((-\infty, x]) = \{w \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$, $\forall x \in \mathbb{R}$, according to Prop. 2.16.

- **To prove**
  $Y = g(X) = \min\{X, 100\}$ is a Borel measurable function, thus, also a r.v., by Corollary 2.40.
- **Proof**
  Firstly, let us remind the reader that a function $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable iff

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

  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 $g^{-1}((-\infty, y]) \in \mathcal{B}(\mathbb{R})$, $\forall y \in I\!R$, according to Remark 2.47 (with $n = 1$).
  Secondly, note that the range of X is

$$\mathbb{R}_X = \mathbb{R}_0^+,$$

  thus, the range of Y is given by

$$\begin{aligned} \mathbb{R}_Y &= g(\mathbb{R}_X) \\ &= [0, 100]. \end{aligned}$$

  Thirdly, $g^{-1}((-\infty, y]) = Y^{-1}((-\infty, y]) = \{\omega \in \Omega : Y(\omega) \leq y\}$ equals:
  - for $y < 0$,
$$\begin{aligned} \{Y \leq y\} &= \emptyset \\ &\in \mathcal{F}; \end{aligned}$$
  - for $0 \leq y \leq 100$,
$$\begin{aligned} \{Y \leq y\} &= \{X \leq y\} \\ &= X^{-1}((-\infty, y]) \\ &\in \mathcal{F}. \end{aligned}$$
  - for $y > 100$,
$$\begin{aligned} \{Y \leq y\} &= \Omega \\ &\in \mathcal{F}. \end{aligned}$$

  Consequently, we can conclude that

$$\{Y \leq y\} \in \mathcal{F}, \forall y \in \mathbb{R},$$

  hence, Y is a Borel measurable function and, thus, a r.v. QED

(b) *Derive and draw the graph of the d.f. of Y and calculate $P(\{Y = 100\})$.* (2.0)

- **D.f. of X**

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\frac{x}{75}}, & x \geq 0 \end{cases}$$

- **D.f. of Y**

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P\left(\min\{X, 100\} \leq y\right) \\ &= \begin{cases} P(\emptyset) = 0, & y < 0 \\ P(X \leq y) = F_X(y) = 1 - e^{-\frac{x}{75}}, & 0 \leq y < 100 \\ P(\Omega) = 1, & y \geq 100 \end{cases} \end{aligned}$$

- **Graph of $F_Y(y)$**

(It has a discontinuity at $y = 100$, etc.)

- **Requested probability**

$F_Y(y)$ has a discontinuity point at $y = 100$, thus,

$$\begin{aligned} P(Y = 100) &= F_Y(100) - F_Y(100^-) \\ &= 1 - (1 - e^{-\frac{100}{75}}) \\ &= e^{-\frac{4}{3}}. \end{aligned}$$

This probability coincides with $P(X > 100)$.

(c) *Describe a method to generate (pseudo-)random numbers from the distribution of Y.* (1.5)

- **Quantile function of X**

Let $u \in (0, 1)$ then the quantile function is given by:

$$\begin{aligned} F_X(x) &= u \\ F_X^{-1}(u) &= x \\ &= -75\ln(1 - u). \end{aligned}$$

- **Quantile transformation**

By Prop. 2.140, if

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

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

$$-75\ln(1 - U) \stackrel{d}{=} X.$$

Consequently, if we want to generate (pseudo-)random numbers from the distribution of $Y = \min\{X, 100\}$ then we have to:

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

(2) assign $x = -75\ln(1 - u)$ and $y = \min\{x, 100\}$.

Group IV — Independence and random variables **6.5 points**

1. *Some form of prophilaxis is said to be 90% effective at prevention during one year's treatment.* (1.5)

   If the degrees of effectiveness in different years are independent, show that the treatment is more likely to fail than not to fail within 7 years.

   - **R.v.**

     $X_i =$ indicator of the effectiveness of treatment i, $i = 1, \ldots, 7$

     $X_i \stackrel{i.i.d.}{\sim} X$, $i = 1, \ldots, 7$

     $X \sim$ Bernoulli$(p = 0.9)$

   - **Requested probability**

$$\begin{aligned} P(\text{all 7 treatments do not fail}) &= P(X_i = 1, \, i = 1, \ldots, 7) \\ &\overset{X^{i.i.d.}\sim X}{=} 0.9^7 \\ &\simeq 0.478297 \\ &< \frac{1}{2}. \end{aligned}$$

   Thus the treatment is more likely to fail than not to fail within 7 years. QED

2. *Each front tire on a particular type of vehicle is supposed to be filled to a pressure of 2.6 psi.*[3]

   Suppose that the actual air pressure in each front tire is a r.v. — X (resp. Y) for the right (resp. left) tire — and that the joint p.d.f. of (X, Y) equals

$$f_{X,Y}(x, y) = \begin{cases} \frac{4}{25}\,xy, & 2 \leq x \leq 3, \, 2 \leq y \leq 3 \\ 0, & \text{otherwise.} \end{cases}$$

   (a) *Prove that X and Y are independent r.v.* (1.0)

   - **Random vector** (X, Y)

     $X =$ air pressure in the left front tire

     $Y =$ air pressure in the right front tire

   - **Joint p.d.f.**

$$f_{X,Y}(x, y) = \begin{cases} \frac{4}{25}\,xy, & 2 \leq x \leq 3, \, 2 \leq y \leq 3 \\ 0, & \text{otherwise} \end{cases}$$

[3] *Non-SI measures of pressure such as pounds per square inch (psi) are used primarily in the USA. 1 psi approximately equals 6894.757 Pa, where pascal (Pa) is the SI unit of pressure (http://en.wikipedia.org/wiki/Pound-force_per_square_inch).*

- **Marginal p.d.f.**

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy \\
&= \begin{cases} \int_2^3 \frac{4}{25}\,xy\,dy = \frac{2x}{25}\left(y^2\big|_2^3\right) = \frac{2x}{5}, & 2 \leq x \leq 3 \\ 0, & \text{otherwise} \end{cases} \\
f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx \\
&= \begin{cases} \int_2^3 \frac{4}{25}\,xy\,dx = \frac{2y}{25}\left(x^2\big|_2^3\right) = \frac{2y}{5}, & 2 \leq y \leq 3 \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

- **Checking if X and Y are independent r.v.**

  Since

$$
\begin{aligned}
f_X(x) \times f_Y(y) &= \begin{cases} \frac{4}{25}\,xy, & 2 \leq x \leq 3,\ 2 \leq y \leq 3 \\ 0, & \text{otherwise} \end{cases} \\
&\equiv f_{X,Y}(x,y),
\end{aligned}
$$

  we can conclude that $X \perp\!\!\!\perp Y$, by Th. 3.38.

(b) *Derive the joint d.f. of (X, Y).* (1.0)

- **Joint d.f. of (X, Y)**

  From the independence criterion in Th. 3.30, we can add that

$$F_{X,Y}(x,y) = F_X(x) \times F_Y(y),\ (x,y) \in \mathbb{R}^2,$$

  where:

$$
\begin{aligned}
F_X(x) &= \int_{-\infty}^{x} f_X(t)\,dt \\
&= \begin{cases} 0, & x < 2 \\ \int_2^x \frac{2t}{5}\,dt = \frac{x^2-4}{5}, & 2 \leq x \leq 3 \\ 1, & x > 3; \end{cases} \\
F_Y(y) &= \begin{cases} 0, & y < 2 \\ \frac{y^2-4}{5}, & 2 \leq y \leq 3 \\ 1, & y > 3. \end{cases}
\end{aligned}
$$

  Hence,

$$
F_{X,Y}(x,y) = \begin{cases} 0, & x < 2 \text{ or } y < 2 \\ \frac{x^2-4}{5} \times \frac{y^2-4}{5}, & 2 \leq x \leq 3 \text{ and } 2 \leq y \leq 3 \\ \frac{x^2-4}{5}, & 2 \leq x \leq 3 \text{ and } y > 3 \\ \frac{y^2-4}{5}, & 2 \leq y \leq 3 \text{ and } x > 3 \\ 1, & x > 3 \text{ and } y > 3. \end{cases}
$$

(c) *Determine the probability that the absolute value of the difference in air pressure between the two front tires is at most 0.2 psi.* (3.0)

- **Requested probability**

$$
\begin{aligned}
P(|Y - X| \leq 0.2) &= P(-0.2 \leq Y - X \leq 0.2) \\
&= P(X - 0.2 \leq Y \leq X + 0.2) \\
&= \int\int_{\{(x,y):\, x-0.2 \leq y \leq x+0.2\}} f_{X,Y}(x,y)\,dy\,dx \\
&= \int_2^{2.2}\int_2^{x+0.2} \frac{4xy}{25}\,dy\,dx \\
&\quad + \int_{2.2}^{2.8}\int_{x-0.2}^{x+0.2} \frac{4xy}{25}\,dy\,dx \\
&\quad + \int_{2.8}^{3}\int_{x-0.2}^{3} \frac{4xy}{25}\,dy\,dx \\
&= \ldots \\
&\overset{Mathematica}{=} 0.363755.
\end{aligned}
$$