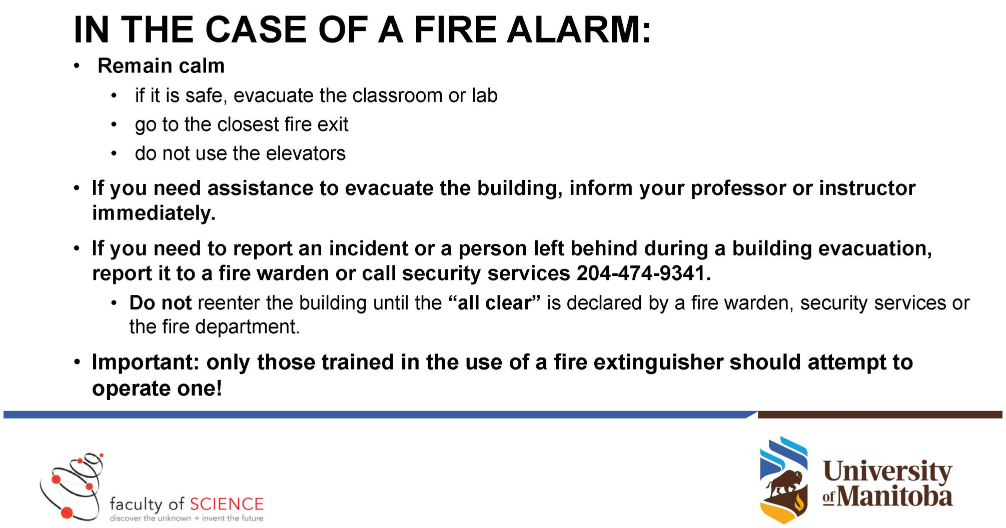
STAT 4100 Lecture Note

Week One (Sep 7 & 9, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Sep/05 12:03:06

Fire Safety Orientation



Contact

- Zhiyang (Gee) Zhou
  - Email: zhiyang.zhou@umanitoba.ca
  - Homepage: zhiyanggeezhou.github.io
  - Office: 330 Machray Hall
- Ke (Kelly) Wang
  - Email: wangk9@myumanitoba.ca
  - Office: TBA

Timeline

- Lectures
  - Mon/Wed/Fri 13:30 – 14:20 at St. Paul's College 258
- Labs
  - Wed 16:30 – 18:00 at Buller 315
- Office Hours
  - (instructor) Mon 14:30 — 15:30 (or by appointment) at 330 Macray Hall
  - (TA) TBD

Grading

- Assignment 20%
  - Submitted via Crowdmark
  - Attaching sourse codes if R is used in computation
  - Always including necessary interpretation
- Midterm 30%
  - In the week of Oct 17
  - Open-book, 3-hour, and online
- Final 50%
  - Open-book and in-person (?)

Meterials

- References
  - (CB) Casella & Berger. 2002. *Statistical Inference*, 2nd Ed.
    - 2 hardcopies reserved at Jim Peebles Science and Technology Library
  - (HMC) Hogg, Mckean & Craig. 2018. *Introduction to Mathematical Statistics*, 8th Ed.
    - Hardcopy of 6th Ed. available
- Lecture notes
  - zhiyanggeezhou.github.io
  - UM Learn
  - Subject to change from time to time without prior notice

Outline

- Topics to be covered
  - Prerequisite
  - Estimation (finite/large sample, optimality)
  - Confidence interval (finite/large sample, interpretation)
  - Hypothesis testing (finite/large sample, optimality, interpretation)

"All models are wrong, but some are useful."

— George Box, *Journal of American Statistical Association* 1976

Statistical modelling & inference

- What are statistical models?
  - Distributions of random variables (r.v.s) of interest
- Statistical inference
  - To answer questions on the underlying statistical models, e.g.,
    - What the model is?
    - Is the r.v. distributed as $N(0, 1)$?

Definitions on random variables (CB/HMC Chp. 1)

- Definition of r.v.: a real-valued function defined on a sample space Ω, i.e.,

$$X = X(\omega), \quad \omega \in \Omega$$

- Cumulative distribution function (cdf) of r.v. X

$$F_X(x) = \Pr(X \leq x)$$

  - Right continuous
  - Non-decreasing
  - Ranging from 0 to 1, i.e., $F_X(-\infty) = 0$ and $F_X(\infty) = 1$

Exercise Lec1.1

- Given $p \in (0, 1)$, suppose

$$F(x) = \begin{cases} 1 - (1-p)^{\lfloor x \rfloor}, & x \geq 1, \\ 0, & \text{otherwise}, \end{cases}$$

  where $\lfloor x \rfloor$ represents the integer part of x. Show that F is a cdf.

- Hint: Check the right-continuity of F at positive integers.

Definitions on random variables (con'd)

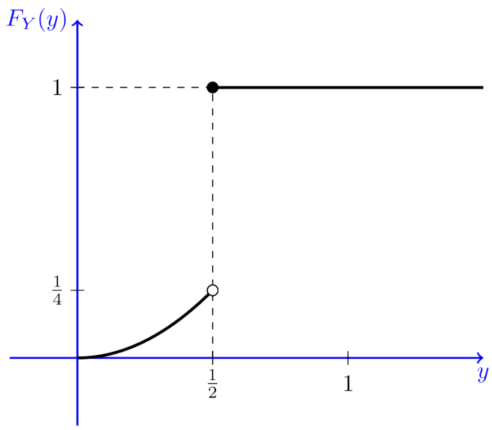
- X is a discrete r.v.
  - X takes countably many values
  - probability mass function (pmf): $p_X(x) = \Pr(X = x)$
- X is a continuous r.v.
  - cdf F_X is absolutely continuous, i.e., $\exists f_X$, s.t.

$$F_X(x) = \int_{-\infty}^{x} f_X(z)\mathrm{d}z, \quad \forall x \in \mathbb{R}.$$

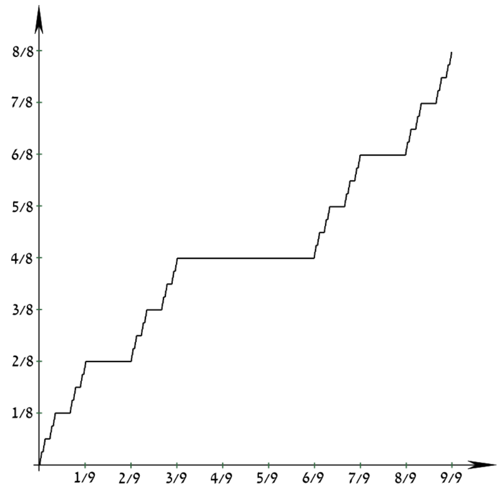
  - f_X is the probability density function (pdf) of X
    * $f_X(x) = (\mathrm{d}/\mathrm{d}x)F_X(x)$ if f_X is continuous at $x \in \mathbb{R}$

- Neither discrete nor continuous
  - X is a mixed r.v., e.g.,

$$F_X(x) = \begin{cases} 1, & x \geq 1/2; \\ x^2, & 0 \leq x < 1/2; \\ 0, & \text{otherwise}. \end{cases}$$

    * 

  - X is following the Cantor distribution



*

Univariate transformation (CB Sec. 2.1 & 2.4)

Support (CB pp. 50)

- In general, for real-valued function g
  - $\text{supp}(g) = \{x \in \text{domain}(g) : g(x) \neq 0\} \subset \text{domain}(g)$
- For discrete r.v. X
  - pmf $p_X(\cdot)$
  - $\text{supp}(X) = \text{supp}(p_X) = \{x \in \mathbb{R} : p_X(x) > 0\}$
  - e.g., Support of $Binom(n, p)$ is $\{0, \ldots, n\}$
- For continuous r.v. X
  - pdf $f_X(\cdot)$
  - $\text{supp}(X) = \text{supp}(f_X) = \{x \in \mathbb{R} : f_X(x) > 0\}$
  - e.g., Support of $N(0, 1)$ is $\mathbb{R}$

Indicator function

- $I_A(x) = 1$ if $x \in A$ and zero otherwise

Find pmf of $Y = g(X)$ given the pmf of X

1. Figure out $\text{supp}(Y) = \{y : y = g(x), x \in \text{supp}(X)\}$
2. Calculate $p_Y(y) = \Pr(Y = y) = \Pr(X \in \{x \in \text{supp}(X) : g(x) = y\})$

Example Lec2.1

Let X have the pmf $p_X(x) = 2^x I_{\{-1,-2,\ldots\}}(x)$. Find the pmf of $Y = X^4$.

Find cdf of $Y = g(X)$ given the distribution of X

- Calculate $F_Y(y) = \Pr\{g(X) \leq y\} = \Pr[X \in g^{-1}\{(-\infty, y]\}]$
  - $g^{-1}\{(-\infty, y]\} = \{x : g(x) \leq y\}$

Example Lec2.2

Let X have the uniform pdf $f_X(x) = \pi^{-1} I_{(-\pi/2,\pi/2)}(x)$. Find the cdf of $Y = \tan X$.

Find pdf of $Y = g(X)$ given the distribution of X

1. Figure out $\text{supp}(Y) = \{y : y = g(x), x \in \text{supp}(X)\}$
2. (Generically) If the cdf F_Y is known OR pdf f_X is easy to be integrated, then

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \int_{\{x:g(x)\leq y\}} f_X(x)\mathrm{d}x$$

- The integration of f_X is often avoidable by employing the Leibniz Rule (CB Thm. 2.4.1):

$$\frac{\mathrm{d}}{\mathrm{d}y} \int_{a(y)}^{b(y)} f(x)\mathrm{d}x = f\{b(y)\}\frac{\mathrm{d}}{\mathrm{d}y}b(y) - f\{a(y)\}\frac{\mathrm{d}}{\mathrm{d}y}a(y)$$

  with $a(y)$ and $b(y)$ both differentiable with respect to y.

---

2. (Alternatively) According to CB Ex. 2.7(b), i.e., an extension of CB Thm. 2.1.5 & 2.1.8 and HMC Thm 1.7.1.

$$f_Y(y) = \sum_{k=1}^{k} f_X\{g_k^{-1}(y)\} \left| \frac{\mathrm{d}}{\mathrm{d}y} g_k^{-1}(y) \right| I_{B_k}(y)$$

- g_k is strictly monotonic on A_k and $g(x) = g_k(x)$ for all $x \in A_k$
  - $\{A_1, \ldots, A_K\}$ is a partition of $\text{supp}(X)$, i.e., $A_k \cap A_{k'} = \emptyset$ if $k \neq k'$ and $\bigcap_{k=1}^{K} A_k = \text{supp}(X)$
- g_k^{-1} is continuously differentiable on $B_k = \{g_k(x) : x \in A_k\}$
- Jacobian (determinant) of g_k^{-1}: $(\mathrm{d}/\mathrm{d}y)g_k^{-1}(y)$

Example Lec2.2'

Let X have the uniform pdf $f_X(x) = \pi^{-1}I_{(-\pi/2,\pi/2)}(x)$. Find the pdf of $Y = \tan X$.

Example Lec2.3

$X \sim \text{Weibull}(\text{shape} = \alpha, \text{scale} = \beta)$, viz. $f_X(x) = (\alpha/\beta)(x/\beta)^{\alpha-1}\exp\{-(x/\beta)^{\alpha}\}I_{(0,\infty)}(x)$. Find the pdf of $Y = \ln(X)$.

Example Lec2.4

Let X have the pdf $f_X(x) = 2^{-1}I_{(0,2)}(x)$. Find the pdf of $Y = X^2$.

Example Lec2.5

Let $f_X(x) = 3^{-1}I_{(-1,2)}(x)$. Find the pdf of $Y = X^2$.