Lecture 25 - 10/21

Friday, October 21, 2022 9:05 AM

Review.

① The distribution of a function of a random variable

Suppose that $Y = g(X)$. X is continuous type. We want to find distribution of Y.

Step 1: identify type of Y (continuous-type or discrete-type) and support of Y (u s.t. $f_Y(u) > 0$)

→ if Y is continuous-type: Step 2: find its cdf. $F_Y(c) = P(Y \leq c) = P(g(X) \leq c) = \int_{u:\, g(u) \leq c} f_X(u)\,du$

Step 3: take derivative of $F_Y(c)$ to derive pdf of Y:

$$f_Y(c) = \frac{\delta F_Y(c)}{\delta c}$$

→ if Y is discrete-type: Step 2: calculate pmf of Y: $P(Y = k) = P(g(X) = k) = \int_{u:\, g(u) = k} f_X(u)\,du$.

② Important example:

X is continuous-type with CDF F_X. $Y = F_X(X)$ is uniformly distributed between 0 and 1.

Today

① Generating random variable with a specified distribution

② Binary hypothesis testing with continuous type observation

---

① Generating random variable with a specified distribution

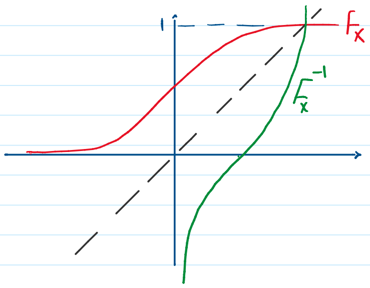
Recall that a random variable X is a function from Ω to $\mathbb{R}$, i.e., $X: \Omega \longrightarrow \mathbb{R}$. For any $\omega \in \Omega$, $X(\omega)$ is a realization or an instance of underlying experiment:

$$\boxed{\text{experiment } (\Omega, \mathcal{F}, P)} \xrightarrow{\text{Generates } \omega} \boxed{\text{Random variable } X} \xrightarrow{\text{Generates } X(\omega)}$$

Recall that we only care about statistical properties, so mapping outcomes of experiment to $\mathbb{R}$ using random variables makes it possible to focus on what matters, i.e., distributions.

We also studied famous distributions that are common in practice. Now what if we want to run a numerical simulation via computer? How should we generate an instance $X(\omega)$, given CDF of X is F_X?

(i). If X is continuous-type, then F_X is increasing, and $Y = F_X(X)$ has $\text{Unif}([0,1])$ distribution. Using this, if U is a uniform r.v. over $[0,1]$, $X = F_X^{-1}(U)$ is a random variable with CDF F_X. Notice that graphically:



F_X^{-1} is resulted by reflection of F via the line passing through origin with slope 1.

Notice that for any $u \in (0,1)$ there exist a unique c_u for which $F_X(c_u) = u$. Hence F_X^{-1} is well defined over $(0,1)$.

$$\boxed{\text{Hence, if } F \text{ is cdf of a continuous-type random variable, and } U \text{ is uniformly distributed over } [0,1] \text{, then } X = F^{-1}(U) \text{ has cdf } F_X = F.}$$

$\Rightarrow$ This suggests to generate a random variable with arbitrary distribution. we can generate a uniform random variable U & pass it through some function g s.t. $X = g(U)$ has the desired distribution.

(ii). Consider a function $F: \mathbb{R} \to [0,1]$ and suppose that F is a cdf, i.e.,

F1. F is increasing;

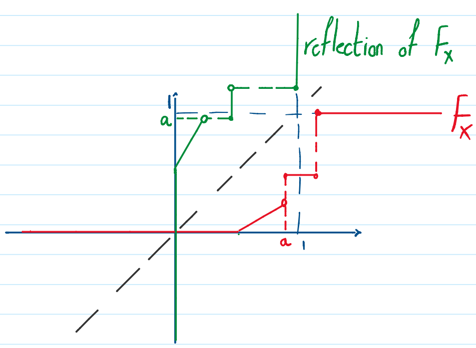
F2. $\lim_{c \to -\infty} F(c) = 0$, $\lim_{c \to +\infty} F(c) = 1$

F3. F is right continuous.

Our goal is to find a function g s.t. the cdf of $X = g(U)$ is given by F, where $U \sim Unif([0,1])$.

We observed that if F is cdf of a continuous-type random variable, then $g = F^{-1}$ is the function we are interested in.

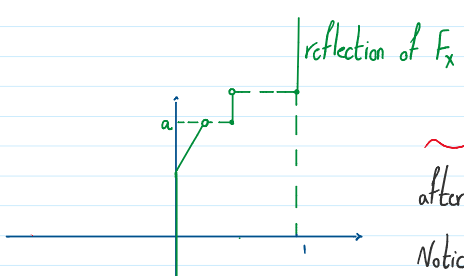
Recall that F^{-1} is reflection of F. Let us naively do the same for general cdf function:



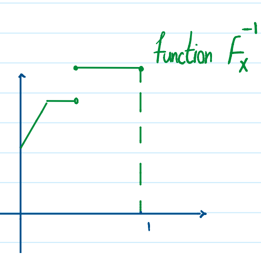
the green plot is reflection of F_X via the line that passes through origin with slop one.

the green lin is not a function, so let's make it into a function with domain $[0,1]$ by

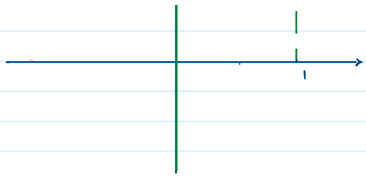
(i) removing vertical line (ii) replacing dashed horizontal lines with solid lines



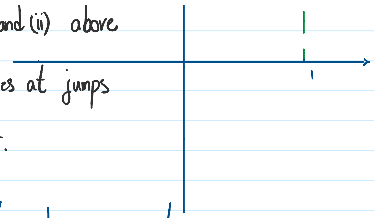
after applying (i) and (ii) above



Notice that values at jumps

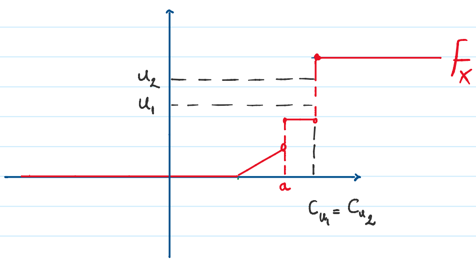


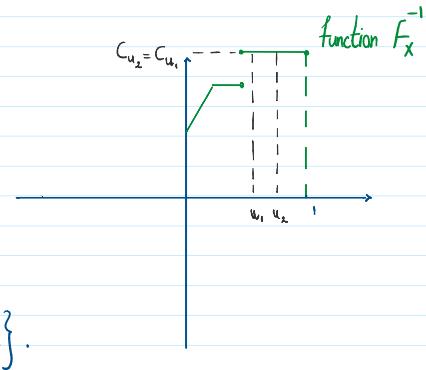
after applying (i) and (ii) above
Notice that values at jumps does not matter.



You can check that the function F^{-1} resulted from the above procedure is

$$F^{-1}(u) = \min\{c : F_X(c) \geq u\} \quad \text{for all } u \in [0,1]$$





Suppose that $c_u = F^{-1}(u) = \min\{c : F(c) \geq u\}$.

We have: $F^{-1}(u) \leq a$ if & only if $c_u \leq a$ iff & only if $u \leq F(a)$. (*)

Hence, if we define $X = F^{-1}(U)$ then cdf of X is given by:

$$F_X(c) = P(X \leq c) = P(F^{-1}(U) \leq c)$$

By (*) $F^{-1}(U) \leq c$ if and only if $U \leq F(c)$.

$$F_X(c) = P(F^{-1}(U) \leq c) = P(U \leq F(c)) = F(c)$$

since for a uniformly distributed random variable U over $[0,1]$, $F_U(a) = P(U \leq a) = a$ if $a \in [0,1]$

So, we have the following statement:

> Suppose that F is a valid CDF. Define $F^{-1}(u) = \min\{c : F(c) \geq u\}$ for any $u \in [0,1]$
> Suppose that U is a uniformly distributed random variable over $[0,1]$. Define $X = F^{-1}(U)$.
> Then cdf of X is F, i.e., $F_X(c) = P(X \leq c) = F(c)$.

Question: How to find F^{-1}?

approach 1: In some cases, F is invertible for values in $(0,1)$, i.e., for any $u \in (0,1)$ there exists a unique c_u s.t. $F(c_u) = u$. In these cases, directly calculate F^{-1}.

approach 2: In some cases graphs are easier to plot. In these cases:

(i) reflect F via the line that passes through origin with slop 1.

(ii) remove vertical lines

(iii) make the resulted graph to a proper functions with domain $[0,1]$ by replacing dashed horizontal lines with solid lines

(iv) notice that values at jump does not matter.

Example:

3.37. **[Generation of random variables with specified probability density function]** Find a function g so that, if U is uniformly distributed over the interval $[0,1]$, and $X = g(U)$, then X has the pdf:

$$f_X(v) = \begin{cases} 2v & \text{if } 0 \le u \le 1 \\ 0 & \text{else.} \end{cases}$$

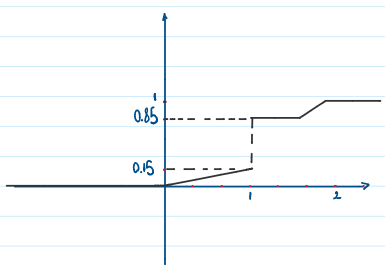
Solution: Notice that $g = F_X^{-1}$. So, the first step is to find F_X.

$$F_X(c) = P(X \leq c) = \begin{cases} 1 & \text{if } c \geq 1 \\ c^2 & \text{if } c \in [0,1] \\ 0 & \text{if } c < 0 \end{cases}$$

Notice that the inverse of F_X, for $u \in (0,1)$ is given by $F_X^{-1}(u) = \sqrt{u}$. Hence the function g is given by $g(u) = \sqrt{u}$ with domain $u \in [0,1]$.

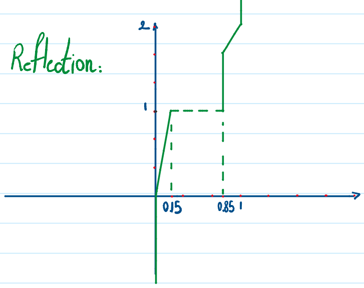
note: value of $g(0)$ and $g(1)$ does not matter since $P(U=0) = P(U=1) = 0$.

Problem. Find a function g s.t. if U is uniformly distributed over $[0,1]$, and $X = g(U)$, then F_X is given by:

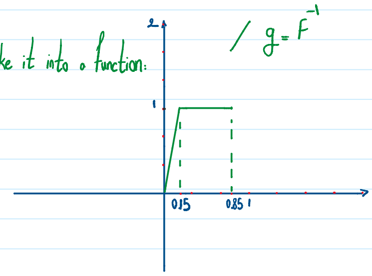


Solution. We reflect F_X via the line that passes through one and then make it into a function.

Reflection:



Make it into a function: / $g = F^{-1}$



② Binary hypothesis testing with continuous type observation

Problem setting is exactly same as the discrete type observation.

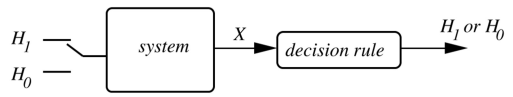
- We have a system that generates continuous-type random variables.
- System is either in state H_0 or H_1
  - if in H_0, pdf of X is given by f_0
  - if in H_1, pdf of X is given by f_1

. if in H_0, pdf of X is given by f_0

. if in H_1, pdf of X is given by f_1

. We observe an outcome from the system, i.e., we observe $\{X=u\}$

. There is a decision rule that assigns hypothesis to each outcome



Based on the decision rule, we decide systen is in state H_0 or H_1.

Notice that $P(X=u)=0$, since X is continuous-type. However, $P\big(X\in(u-\varepsilon,u+\varepsilon)\big)\approx 2\varepsilon f_X(u)$.

So, for continuous-type observations we focus on pdf instead of pmf. Hence, we define the likelihood ratio by

$$\Lambda(u)=\frac{f_1(u)}{f_0(u)}$$

. We are only interested in threshold policies :

(i) general threshold policies with threshold $\tau>0$:

$$\begin{cases} \text{if } \Lambda(u)>\tau \text{ then } H_1 \\ \text{if } \Lambda(u)<\tau \text{ then } H_0 \\ \text{if } \Lambda(u)=\tau, \text{ either} \end{cases}$$

(ii) ML. threshold policy with $\tau=1$

(iii) MAP. threshold policy with $\tau=\frac{\pi_0}{\pi_1}$, $\pi_0=P(\text{system in } H_0)$, $\pi_1=P(\text{system in } H_1)$ are priors.

. Instead of likelihood ration we can consider loglikelihood:

$$\log\Lambda(u)=\log f_1(u)-\log f_0(u)$$

& compare it with $\log\tau$.

note: If u is not in support of f_1 then it means system is in H_0.

If u is not in support of f_0 then it means system is in H_1.

. We have simillar notions as befor:

$$P_{\text{false alarm}}=P(\text{Decision rule is } H_1 \mid \text{System in } H_0)$$

$$P_{\text{miss}}=P(\text{Decision rule is } H_0 \mid \text{System in } H_1)$$

$$P_{\text{error}}=P(\text{Decision rule}\neq\text{state of system})$$

$$=\pi_0 P_{\text{false alarm}}+\pi_1 P_{\text{miss}}$$ ⇝ same as in discrete-type.

Example:

3.31. **[A simple hypothesis testing problem with continuous-type observations]** Consider the hypothesis testing problem in which the pdf's of the observation X under hypotheses H_0 and H_1 are given, respectively, by:

3.31. **[A simple hypothesis testing problem with continuous-type observations]**
Consider the hypothesis testing problem in which the pdf's of the observation X under hypotheses H_0 and H_1 are given, respectively, by:

$$f_0(u) = \begin{cases} \frac{1}{2} & \text{if } -1 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_1(u) = \begin{cases} |u| & \text{if } -1 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Assume the priors on the hypotheses satisfy $\pi_1 = 2\pi_0$.

(a) Find the MAP rule.

(b) Find $p_{\text{false alarm}}$, p_{miss} and the average probability of error, p_e, for the MAP rule.

Solution: Notice that the likelihood ratio is given by $\Lambda(u) = \frac{f_1(u)}{f_0(u)} = 2|u|$ for $u \in [-1,1]$.

(a) We have to compare $\Lambda(u)$ with $\frac{\pi_0}{\pi_1} = \frac{1}{2}$ for $u \in [-1,1]$ (notice that support of f_1 & f_0 are both $[-1,1]$)

The decision rule is:

$$\begin{cases} H_1 & \text{if } 1 \ge |u| > 1/4 \\ H_0 & \text{if } |u| < 1/4 \\ \text{either} & \text{if } |u| = 1/4 \end{cases}$$

(b) $\pi_1 = 2\pi_0$ and $\pi_0 + \pi_1 = 1 \Rightarrow \pi_0 = 1/3$ and $\pi_1 = 2/3$

$P_{\text{false alarm}} = P(\text{Decision rule is } H_1 \mid \text{system in } H_0)$

$$= \int_{u:\ \text{Decision for } u \text{ is } H_1} f_0(u)\,du = \int_{u:\ |u| \in (1/4, 1]} \tfrac{1}{2}\,du = \tfrac{1}{2} \cdot \tfrac{6}{4} = \tfrac{3}{4}$$

$P_{miss} = P(\text{Decision rule is } H_0 \mid \text{system in } H_1)$

$$= \int_{u:\ \text{Decision for } u \text{ is } H_0} f_1(u)\,du = \int_{-1/4}^{1/4} |u|\,du = 2\int_0^{1/4} u\,du = \frac{1}{16}$$

$$P_{error} = \pi_0 P_{\text{false alarm}} + \pi_1 P_{miss} = \frac{1}{3} \cdot \frac{3}{4} + \frac{2}{3} \cdot \frac{1}{16} = \frac{7}{24}$$