Lecture 26 - 10/24

Saturday, October 22, 2022 11:55 AM

Review.

① Distribution of a function of random variable

$Y = g(X)$, pdf of X is given, pdf of Y is desired

Step 1: Find type of Y & its support

Step 2: If continuous-type find its cdf, if discrete type find its pmf

Step 3. If continuous-type find its pdf by differentiating

② Generating variable with specified distribution

. Given $U \sim unif([0,1])$ and a cdf F, find function g such that the cdf of $X = g(U)$ is F, i.e., for any $c \in \mathbb{R}$ $F_X(c) = F(c)$.

. Answer: $g(u) = F^{-1}(u) = \min\{c : F(c) \geq u\}$.

(a) if possible calculate F^{-1}

(b) if it is hard use reflection & then make it into a function.

③ Hypothesis testing for continuous type observation.

(i) The system generates continuous-type random variable X

. If the system is in state H_0, the pdf of X is given by f_0

. If the system is in state H_1, the pdf of X is given by f_1

(ii) We observe a realization of X, i.e, we observe $X = u$.

(iii) We guess the state of system, using our observation, based on the decision rule.

We are only interested in threshold decision rule that compare $\Lambda(u) = \frac{f_1(u)}{f_0(u)}$ with threshold τ.

ML: $\tau = 1$, MAP: $\tau = \frac{\pi_0}{\pi_1}$

Today: ① quick recap of cdf

② Joint cdf

③ Joint pmf

---

① quick recap of cdf

Recall that cdf of a random variable X is defined as $F_X(c) = P(X \leq c)$, for any $c \in \mathbb{R}$. Also recall that a function F is a cdf if and only if it satisfies the following:

F.1. F is increasing

F.2: $\lim_{c \to \infty} F(c) = 1$, $\lim_{c \to -\infty} F(c) = 0$

F.3. F is right-continuous, i.e., $\lim_{a \to c^+} F(a) = F(c)$

cdf of a random variable uniquely determines probability of any event $P(X \in A)$ for $A \subseteq \mathbb{R}$. (A measurable!)

We were interested in F since we were interested in distribution of X. What if we are interested in distribution of

cdf of a random variable uniquely determines probability of any event $P(X\in A)$ for $A\subseteq\mathbb{R}$. (A measurable!)
We were interested in F_X, since we were interested in distribution of X. What if we are interested in distribution of multiple random variable at the same time?

② joint cdf

Def. Suppose that X and Y are two random variables defined over the same probability space $(\Omega,\mathcal{F},P)$, i.e., $X:\Omega\to\mathbb{R}$ and $Y:\Omega\to\mathbb{R}$
The joint cumulative distribution function (joint cdf) is the function of two variables defined as

$$F_{X,Y}(u_o,v_o) = P\{X\leq u_o,\ Y\leq v_o\} \quad \text{for any } u_o\in\mathbb{R} \text{ and } v_o\in\mathbb{R}.$$

example: Suppose that X denote the temprature of your mobile while gaming, & Y denotes its CPU usage. We are interested in probability of the event $\{w\in\Omega:\ X(w)\leq 50F^\circ \text{ and } Y(w)\leq 45\%\}$

$$F_{X,Y}(50,0.45) = P(X\leq 50,\ Y\leq \tfrac{45}{100}) = P(\{w\in\Omega:\ X(w)\leq 50F^\circ \text{ and } Y(w)\leq 45\%\})$$

Notice that X & Y are defined over the same space.
Notice that the event $\{X\leq u_o,\ Y\leq v_o\}$ means (X,Y) is in the shaded area below:

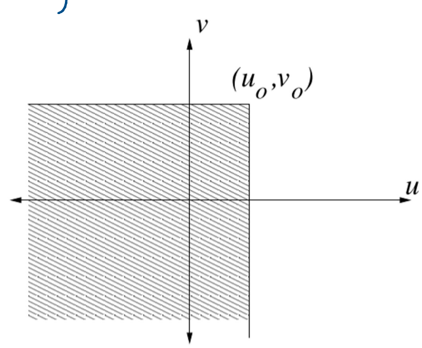

Figure 4.1: Region defining $F_{X,Y}(u_o,v_o)$.

Joint cdf uniquely determines the distribution of any joint event $P((X,Y)\in A)$ where $A\subseteq\mathbb{R}^2$. (A measurable!)
In particular, if $A=(a,b]\times(c,d]$ we have

$$P((X,Y)\in A) = P(a<X\leq b,\ c<Y\leq d) = F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c)$$

$$= P(X\leq b, Y\leq d) - P(X\leq b, Y\leq c) - P(X\leq a, Y\leq d) + P(X\leq a, Y\leq c)$$

we have deducted this area twice so we have to add one back
=> sign is +

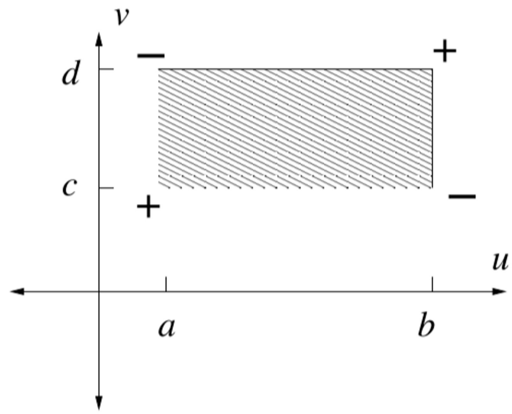

Figure 4.2: $P\{(X,Y) \in \text{shaded region}\}$ is equal to $F_{X,Y}$ evaluated at the corners with signs shown.

Joint cdf and marginal cdf are related to each other.

Proposition 4.1.1. Suppose that $F_{X,Y}$ is joint cdf of X and Y.

- $\lim_{v\to\infty} F_{X,Y}(u,v) = \lim_{v\to\infty} P(X\leq u, Y\leq v) = P(X\leq u) = F_X(u)$ ⟶ called marginal cdf of X
- $\lim_{u\to\infty} F_{X,Y}(u,v) = \lim_{u\to\infty} P(X\leq u, Y\leq v) = P(Y\leq v) = F_Y(u)$ ⟶ called marginal cdf of Y

proof: Let $G_0 = \{X\leq u, Y\leq 0\}$ and define $G_n = \{X\leq v, n-1 < Y \leq n\}$ for any $n\geq 1$.

$$F_X(u) = P(X\leq u) = P(G_0 \cup G_1 \cup G_2 \cup \cdots)$$
$$= \lim_{n\to\infty} P(G_0) + \cdots + P(G_n)$$
$$= \lim_{n\to\infty} P(G_0 \cup \cdots \cup G_n)$$
$$= \lim_{n\to\infty} P(X\leq u, Y\leq n) = \lim_{n\to\infty} F_{X,Y}(u,n)$$

Since $F_{X,Y}(u,v)$ is non-decreasing in v, $\lim_{v\to\infty} F_{X,Y}(u,v)$ is well-defined, and hence equals $F_X(u)$.

Similar to cdf, joint cdf also satisfies some properties. Moreover any function that satisfies these is a joint cdf.

Proposition: A function F is a joint cdf of some pair of random variables if and only if:

JF.1: $0 \leq F(u,v) \leq 1$ for any $(u,v)\in\mathbb{R}^2$

JF.2: $F(u,v)$ is nondecreasing in u and nondecreasing in v.

JF.3: $F(u,v)$ is right-continuous in u and right-continuous in v.

JF.4: $\lim_{u\to-\infty} F(u,v) = 0$ and $\lim_{v\to-\infty} F(u,v) = 0$

JF.5: $\lim_{u\to\infty}\lim_{v\to\infty} F(u,v) = 1$

$\boxed{\text{JF.6. For any } a<b \text{ and } c<d,\ F(b,d) - F(a,d) - F(b,c) + F(a,c) \geq 0}$ ⟶ This is due dimension! being non-decreasing won't guarantee this.

③ Joint pmf

Def. Suppose that X and Y are two discrete-type random variables defined over the same probability space $(\Omega, \mathcal{F}, P)$. The joint probability mass function (joint pmf) is denoted by $p_{X,Y}$ and definned as:

$$p_{X,Y}(u,v) = P(X=u, Y=v)$$

Remark: Since X and Y are discrete type, they can only take at most countably many values:

- there exists $u_1, u_2, \dots$ s.t. $P(X \in \{u_1, u_2, u_3, \dots\}) = 1$
- there exists $v_1, v_2, \dots$ s.t. $P(Y \in \{v_1, v_2, \dots\}) = 1$

Joint pmf and marginal pmfs are related:

- $p_X(u) = P(X=u) = P(X=u, Y \in \{v_1, v_2, \dots\}) = \sum_i P(X=u, Y=v_i) = \sum_i p_{X,Y}(u, v_i)$ ⇝ marginal pmf of X
- $p_Y(v) = P(Y=v) = P(Y=v, X \in \{u_1, u_2, \dots\}) = \sum_i P(Y=v, X=u_i) = \sum_i p_{X,Y}(u_i, v)$ ⇝ marginal pmf of Y

p_X and p_Y are called marginal pmf.