Math 621 Lec 5 9/12/17

$\vec{X}$ is a vector of r.v.'s s.t. $\dim[\vec{X}] = k$.

$$\vec{\mu} := E[\vec{X}] = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_k] \end{bmatrix}$$

$$\Sigma := Var[\vec{X}] = \begin{bmatrix} Var[X_1] & Cov[X_1, X_2] & \cdots \\ Cov[X_2, X_1] & Var[X_2] & \ddots \\ \vdots & & \ddots \\ & & & Var[X_k] \end{bmatrix} = \{Cov[X_i, X_j]\}_{\substack{i=1\ldots k \\ j=1\ldots k}}$$

$$\Sigma_0 := Corr[\vec{X}] = \begin{bmatrix} 1 & Corr[X_1, X_2] & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{bmatrix} = \{Corr[X_i, X_j]\}$$

let $T = X_1 + \ldots + X_k = \vec{1}^T \vec{X} = [1 \cdots 1] \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}$

$$E[T] = \sum_{i=1}^{k} \mu_i = \vec{1}^T \vec{\mu}$$

$$Var[T] = Var[\vec{1}^T \vec{X}] = Cov[\Sigma X_i, \Sigma X_i] = \sum_{j=1}^{k} \sum_{i=1}^{k} Cov[X_i, X_j]$$

let $Y = \vec{c}^T \vec{X}$

$$E[Y] = \Sigma c_i \mu_i = \vec{c}^T \vec{\mu}$$

$$Var[Y] = Var[\vec{c}^T \vec{X}] = ?$$

If $A \in \mathbb{R}^{n \times n}$, $\vec{c} \in \mathbb{R}^n$

What is $\vec{c}^\top A \vec{c}$? Called a "quadratic form"...

$$= \vec{c}^\top \begin{bmatrix} c_1 a_{11} + \ldots + c_n a_{1n} \\ c_1 a_{21} + \ldots + c_n a_{2n} \\ \vdots \\ c_1 a_{n1} + \ldots + c_n a_{nn} \end{bmatrix} = \begin{array}{l} c_1^2 a_{11} + c_1 c_2 a_{12} + \ldots + c_1 c_n a_{1n} + \\ c_2 c_1 a_{21} + c_2^2 a_{22} + \ldots + c_2 c_n a_{2n} + \\ \vdots \\ c_n c_1 a_{n1} + c_n c_2 a_{n2} + \ldots + c_n^2 a_{nn} \end{array}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} c_i c_j a_{ij}$$

a $1_{a} \ldots$

What is $\text{Var}\left[\vec{c}^\top \vec{X}\right]$?

$$= \text{Var}\left[c_1 X_1 + \ldots + c_K X_K\right]$$

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} \text{Cov}\left[c_i X_i, c_j X_j\right]$$

elements of the Var-Cov matrix

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} c_i c_j \text{Cov}\left[X_i, X_j\right]$$

$$= \vec{c}^\top \text{Var}\left[\vec{X}\right] \vec{c}$$

Markowitz Portfolio theory

$X_1, \ldots, X_K$ ~~are~~ are the returns of K assets

$w_1, \ldots, w_K$ are the allocation weights s.t. $\vec{1}^\top \vec{w} = 1$

$V = \vec{w}^\top \vec{X} \Rightarrow \mu_V = \vec{w}^\top \vec{\mu}$

$\sigma_V^2 = w^\top \text{Var}(\vec{X})\, w$

$\min \sigma_V^2$ s.t. $\mu_V = \mu_0$ and $\vec{1}^\top \vec{w} = 1$

we will solve this lecture

$\vec{X} \sim \text{Multinomial}(n, \vec{p})$, $\quad E[\vec{X}] = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_K] \end{bmatrix} = \begin{bmatrix} np_1 \\ \vdots \\ np_K \end{bmatrix} = n\vec{p}$

Makes sense since

$X_1 \sim \text{Bin}(n, p_1), \ldots, X_K \sim \text{Bin}(n, p_K)$ but not independent!

$\text{Var}\left[\vec{X}\right] = ?$

What is $\text{Var}\left[X_1\right]$? $\quad = np_1(1-p_1) \quad$ See notes from prob. class

$$Var[\vec{X}] = \begin{bmatrix} np_1(1-p_1) & Cov[X_1, X_2] & & \\ Cov[X_2, X_1] & np_2(1-p_2) & & \\ \vdots & & \ddots & \\ \vdots & & & np_K(1-p_K) \end{bmatrix}$$

Need $Cov[X_i, X_j]$

1st strategy

$$= E[X_i X_j] - \mu_i \mu_j$$

$$= \left(\sum_{x_i \in Supp[X_i]} \sum_{x_j \in Supp[X_j]} x_i x_j \, \underbrace{p(x_i, x_j)} \right) - np_i(1-p_i)\, np_j(1-p_j)$$

Difficult to get!

this is the marginal prob of X_1, X_2. Difficult... maybe on HW!

Next strategy...

Recall if

$$X \sim Bin(n, p) \iff X = \sum_{i=1}^{n} X_i \quad \text{where } X_1, \ldots, X_n \overset{iid}{\sim} Bern(p) = Bin(1, p)$$

if

since all marginals are binomial....

$$\begin{bmatrix} X_1 \\ \vdots \\ X_K \end{bmatrix} \sim Multin(n, \vec{p}) \iff X_1 = \sum_{i=1}^{n} X_{i1} \quad X_{11}, \ldots, X_{n1} \overset{iid}{\sim} Bern(p_1)$$

$$\vdots$$

$$X_K = \sum_{i=1}^{n} X_{iK} \quad \text{where } X_{1K}, \ldots, X_{nK} \overset{iid}{\sim} Bern(p_K)$$

but X_{i1}, X_{iK} are not indep!

Further $\vec{X} \sim \text{Multinom}(n, \vec{p}) \Rightarrow \vec{X} = \sum_{i=1}^{n} \overset{\text{col.}}{\vec{X}_{i\cdot}}$ s.t. $\vec{X}_{1\cdot}, \ldots, \vec{X}_{n\cdot} \overset{iid}{\sim} \text{Multinom}(1, \vec{p})$

Was on the HW!

NOW... $Cov[X_i, X_j] = Cov\left[\sum_{l=1}^{n} X_{li}, \sum_{h=1}^{n} X_{hj}\right]$

indep. across cols, dep. across rows

$$\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{12} & X_{22} & & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{1K} & X_{2K} & \cdots & X_{Kn} \end{bmatrix}$$

(All Bernoullis)

$$= \sum_{l=1}^{n} \sum_{h=1}^{n} Cov[X_{li}, X_{hj}]$$

$$= \sum_{l=1}^{n} \sum_{h=1}^{n} E[X_{li} X_{hj}] - p_i p_j$$

if $l \neq h$, this is a different multinom model so $E[X_{li} X_{hj}] = E[X_{li}]\,E[X_{hj}] = p_i p_j$

thus all these terms are zero. Now... $l = h$, what happens?

$$= \sum_{l=1}^{n} E[X_{li} X_{lj}] - p_i p_j$$

$E[X_{li} X_{lj}] = 0$

$p(x_i, x_j)$

X_{lj} \ X_{li}	0	1
0	+	+
1	+	0

not possible! Mutually exclusive

$X_{li} X_{lj}$

X_{lj} \ X_{li}	0	1
0	0	0
1	0	1

How many terms where $l = h$? n...

$$\Rightarrow Cov[X_i, X_j] = \overbrace{-n p_i p_j}^{\text{actually negative}}$$

$\uparrow \; \uparrow \; \uparrow$

$+ \; + \; +$

Makes sense? Yes if #C's $\uparrow$ $\Rightarrow$ #A's $\downarrow$ since there's a fixed # of trials.

HW: $Corr[X_i, X_j]$, matrix notation

Continuous r.v.'s have prob. density functions (PDFs)

$p(x) = 0 \ \forall x$. Why? $\nexists\, p$ s.t. $\sum_{x \in S} p(x) < \infty$ s.t. S is uncountable

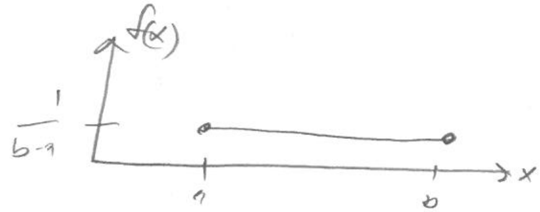
$$f(x) := F'(x) \quad \text{where} \quad F(x) \text{ is the CDF}$$

The domain of f is the $Supp(X) := \{x : f(x) > 0\}$ o/t $f(x) = 0$

and $|Supp(X)| = |\mathbb{R}|$ i.e. uncountabl. infinite

A basic continuous r.v. is the uniform r.v.

$X \sim U(a,b) := \frac{1}{b-a}$

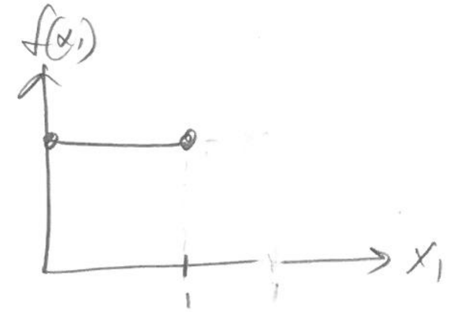


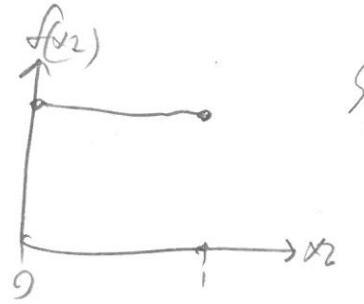
Parameter space: $a, b \in \mathbb{R}$ s.t. $a < b$

$Supp(X) = [a,b]$

the "standard" uniform is $a=0, b=1 \Rightarrow X \sim U(0,1) := 1$

What does $T_2 := X_1 + X_2 \sim$? If $X_1, X_2 \overset{iid}{\sim} U(0,1)$?





$Supp(T_2) = [0, 2]$

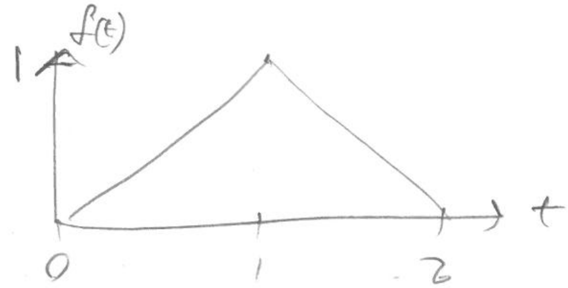
Intuitively, how often could $T_2 = 0$? well $X_1 = 0$ & $X_2 = 0 \Rightarrow$ rare common

'' '' '' $T_2 = 2$? well $X_1 = 1$ & $X_2 = 1 \Rightarrow$ rare $\Uparrow$

'' '' '' $T_2 = 1$? $X_1 = 0$ & $X_2 = 1$, $X_2 = 1$ & $X_1 = 0$, $X_1 = 0.5$ & $X_2 = 0.5$, $X_1 = \frac{1}{4}$ & $X_2 = \frac{3}{4}$, etc...

$T_2 = 1.9$

$X_1 = 1$ & $X_2 = 0.9$, $X_1 = 0.95$ & $X_2 = 0.95$, $X_1 = 0.5$? No... $\Rightarrow$ Rare



It seems this should be the triangular distr.

Proof

$$f_{T_2}(t) = \int_{x \in Supp(X_1)} f_{X_1}(x)\, f_{X_2}(t-x)\, dx \qquad x - t \in [-1,0] \Rightarrow x \in [t-1,t]$$

$f_{X_1}(x) * f_{X_2}(x) ::=$

$$= \int_0^1 1\, \overbrace{\mathbb{1}_{x\in[0,1]}}^{\text{not needed}}\; 1\, \mathbb{1}_{t-x\in[0,1]}\, dx$$

Convolutions for densities... Same concept as before...

$$= \int_0^1 \mathbb{1}_{x\in(t-1,t]}\, dx$$

$$= \int_{\max\{0,t-1\}}^{\min\{1,t\}} dx = x\Big]_{\max\{0,t-1\}}^{\min\{1,t\}} = \min\{1,t\} - \max\{0,t-1\}$$

this is the answer for $t \in (0,2)$

alternatively... $f_{T_2}(t) = \begin{cases} \text{if } t<1 \Rightarrow t \\ \text{if } t \geq 1 \Rightarrow 1-(t-1) = 2-t \end{cases} \mathbb{1}_{t\in(0,2)}$