Jan B. CW

Math 621 Prob. Aug. 27, 2019

Intro to Distribution Theory

defined as $:=$

A discrete random variable (r.v.) X has probability mass function (PMF) $p(x) := P(X=x)$ notation: $X \sim p(x)$ and cumulative distribution function (CDF) $F(x) := P(X \leq x)$

r.v. X has "support" $Supp(X) = \{x : p(x) > 0, x \in \mathbb{R}\}$

(support is all values of X that have a non-zero probability)

realizations means the possible values of X

Support and PMF are related via $\sum_{x \in Supp(X)} p(x) = 1$

Also $|Supp(X)| \leq |\mathbb{N}|$

i.e. the number of possible different realizations is finite or at most countably infinite

In my opinion, the most fundamental r.v. is Bernoulli:

Bernoulli r.v.: $X \sim Bern(\overset{\text{parameter}}{p}) := \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } 1-p \end{cases}$

also written:

parameter space is all possible values of the parameter

Support: $Supp(X) = \{0,1\}$

$X \sim Bern(p) := p^x(1-p)^{1-x}$

X can only be 0 or 1

$p(1) = p \quad p(0) = 1-p$

p is its parameter (tuning knob) to the "parameter space" $p \in (0,1)$

Degenerate Cases

If $p=1$, $X \sim \text{Bern}(1) = \{1 \text{ wp } 1 = \text{Deg}(1)$
(X is always 1) (degenerate r.v.)

If $p=0$, $X \sim \text{Bern}(0) = \{0 \text{ wp } 1 = \text{Deg}(0)$
(X is always 0)

$X \sim \text{Deg}(c) = \{c \text{ wp } 1$ ← X can only be c
$P(X=c)=1$

Degenerate

$\underbrace{P(X=3.7)}_{=0} = p(3.7) = p^{3.7}(1-p)^{-2.7} = 0.5$

← must be 0 — if $p=\frac{1}{2}$ — did not get 0

so we don't get this nonsense, we use indicator functions so get probability 0 for everything outside the support

is defined
$:=$

let $\mathbb{1}_A := \begin{cases} 1 & \text{if } A \\ 0 & \text{if not } A \end{cases}$ — A is a condition (statement that can be true or false)

indicator of A

indicator function, notation can be
$\mathbb{1}(A)$, $1(A)$, $I(A)$

ex: $f(x) = \mathbb{1}_{x<17}$ so $f(19)=0$
$f(-173)=1$

use: $\forall x \in \mathbb{R}$
$$p(x) = p^x(1-p)^{1-x}\,\mathbb{1}_{x \in \underbrace{\{0,1\}}_{\text{supp}(X)}}$$
for $X \sim \text{Bern}(p)$

$$p(x) = p_{\text{old}}\,\mathbb{1}_{x \in \text{supp}(X)}$$

$X \sim Bern(p) \rightarrow p(x) = p^x (1-p)^{1-x} \mathbb{1}_{x \in \{0,1\}}$

$X_1, X_2, \ldots, X_n$ are discrete r.v.s. which have a joint mass function (JMF)

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

and and and

If $X_1, X_2, \ldots X_n$ are independent

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i) \quad \forall \vec{x} \in \mathbb{R}^n$$

$\vec{x} = (x_1, x_2, \ldots, x_n)$

If $X_1, X_2, \ldots, X_n$ are indep. and identically distributed (i.i.d.)

$X_1, X_2, \ldots, X_n \overset{i.i.d}{\sim}$

$$p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)$$

A convolution:

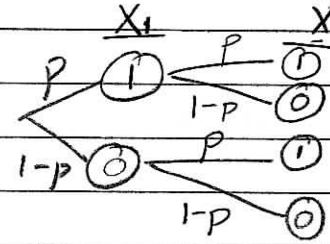
Let $X_1, X_2 \overset{i.i.d.}{\sim} Bern(p)$

Let $T = X_1 + X_2 \sim p_T(t)$

$$Supp[T] = Supp[X_1] + Supp[X_2] = \{0, 1, 2\}$$

where $A + B := \{a + b : a \in A, b \in B\}$

$X_1, X_2 \overset{i.i.d}{\sim} \text{Bern}(p)$ and $T = X_1 + X_2$



$\langle x_1, x_2 \rangle$	p_{X_1,X_2}	T	$p_T(t)$
$\langle 1,1 \rangle$	p^2	2	p
$\langle 1,0 \rangle$	$p(1-p)$	1	$\left.\right\} 2(1-p)$
$\langle 0,1 \rangle$	$(1-p)p$	1	
$\langle 0,0 \rangle$	$(1-p)^2$	0	$(1-p)^2$

so...

$$p_T(t) \begin{cases} 0 & \text{wp } (1-p)^2 \\ 1 & \text{wp } 2p(1-p) \\ 2 & \text{wp } p^2 \end{cases}$$

check if adds up to 1

$$\sum_{t \in supp(T)} p_T(t) = (1-p)^2 + 2p(1-p) + p^2$$
$$= 1 - 2p + p^2 + 2p - 2p^2 + p^2$$

$$P(T=t) = p_T(t) = \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} p_{X_1,X_2}(x_1, x_2) \mathbb{1}_{\substack{t = x_1 + x_2 \\ x_1 \in \{0,1\} \\ x_2 \in \{0,1\}}}$$

use $\sum_{x \in \mathbb{R}} x \mathbb{1}_{x=17} = 17$

only adding up 1 value

$$p_T(t) = \sum_{x_1 \in \mathbb{R}} \sum_{x_2 \in \mathbb{R}} p_{X_1,X_2}(x_1, t - x_1) \mathbb{1}_{x_2 = t - x_1}$$

$$p_T(t) = \sum_{x_1 \in \mathbb{R}} p_{X_1,X_2}(x, t-x)$$

general discrete convolution formula:

$$p_T(t) = \sum_{x \in \mathbb{R}} p_{X_1,X_2}(x, t-x)$$ where $T = X_1 + X_2$, $T \sim p_T(t)$

if X_1, X_2 are independent, we can simplify...

$$p_T(t) = \sum_{x \in \mathbb{R}} p_{X_1}(x) p_{X_2}(t-x)$$

if X_1, X_2 are i.i.d. (both have PMF p)

$$p_T(t) = \sum_{x \in \mathbb{R}} p(x) p(t-x)$$

if X_1, X_2 are i.i.d. (with PMF p) $\qquad T = X_1 + X_2$

$$p_T(t) = \sum_{x \in \mathbb{R}} p(x)\, p(t-x)$$

Our case... $(X_1, X_2 \overset{i.i.d.}{\sim} \text{Bern}(p))$ $\qquad T = X_1 + X_2$

$$p_T(t) = \sum_{x \in \mathbb{R}} \left(p^x (1-p)^{1-x} \mathbb{1}_{x \in \{0,1\}}\right)\left(p^{t-x}(1-p)^{1-(t-x)} \mathbb{1}_{t-x \in \{0,1\}}\right)$$

$$= \sum_{x \in \mathbb{R}} p^x (1-p)^{1-x}\, p^{t-x} (1-p)^{1-t+x}\, \mathbb{1}_{\substack{t-x \in \{0,1\} \\ x \in \{0,1\}}}$$

alternative version found in books:

$$\sum_{x \in \mathbb{R}} p_{X_1}(x)\, p_{X_2}(t-x)$$

$$= \sum_{x \in \mathbb{R}} p(x)\, p(t-x) \qquad \text{if } X_1, X_2 \text{ are i.i.d.}$$

$$= \sum_{x \in \mathbb{R}} p_{OLD} \mathbb{1}_{x \in Supp[X]}\, p_{OLD} \mathbb{1}_{t-x \in supp(X)}$$

$$= \sum_{x \in Supp[X]} p_{OLD}(x)\, p_{OLD}(t-x)\, \mathbb{1}_{t-x \in supp[X]}$$

Our case

$$p_T(x) = \sum_{x\in\mathbb{R}} \left(p^x(1-p)^{1-x}\mathbb{1}_{x\in\{0,1\}}\right)\left(p^{t-x}(1-p)^{1-(t-x)}\mathbb{1}_{t-x\in\{0,1\}}\right)$$

$$= \sum_{x\in\mathbb{R}} p^x(1-p)^{1-x}p^{t-x}(1-p)^{1+x-t}\mathbb{1}_{t-x\in\{0,1\}}$$

$$= p^t(1-p)^{t-x}\sum \mathbb{1}_{t-x\in\{0,1\}}$$

$$= p^t(1-p)^{t-x}\left(\mathbb{1}_{t\in\{0,1\}} + \mathbb{1}_{t-1\in\{0,1\}}\right)$$

$$= p^t(1-p)^{t-x}\left(\mathbb{1}_{t\in\{0,1\}} + \mathbb{1}_{t\in\{1,2\}}\right)$$

$$= \left.\begin{cases} (1-p)^2 & \text{if } t=0 \\ 2p(1-p) & \text{if } t=1 \\ p^2 & \text{if } t=2 \\ 0 & \text{otherwise} \end{cases}\right\} = \binom{2}{t}p^t(1-p)^{2-t}$$

$$\binom{2}{t} = \begin{cases} 1 & \text{if } t=0 \\ 2 & \text{if } t=1 \\ 1 & \text{if } t=2 \\ 0 & \text{otherwise} \end{cases}$$

so

$$p_T(x) = \binom{2}{t}p^t(1-p)^{2-t} = \text{Binom}(2,p)$$

so $T = X_1 + X_2$ is $\text{Binom}(2,p)$ if X_1, X_2 are i.i.d. $\text{Bern}(p)$