

Math 180A: Introduction to Probability

Lecture A00 (Au)

math.ucsd.edu/~bau/w21.180a

Lecture B00 (Nemish)

math.ucsd.edu/~ynemish/teaching/180a

Today: ASV 6.1, 6.2

Video: Prof. Todd Kemp, Fall 2019

Next: ASV 6.2, 6.3

Week 8: Homework 6 (due Wednesday, Feb 24)

Midterm 2 (Wednesday, Feb 24) covers up to and including MGF

Practice midterm posted (solutions will be posted by Monday, Feb 22)

Definitions: Given (discrete) random variables X_1, X_2, \dots, X_n all defined on the same sample space, their joint distribution is the collection of all

$$\left\{ \begin{array}{l} P(X_1=k_1, X_2=k_2, \dots, X_n=k_n) \\ \text{all possible values } k_1 \text{ of } X_1, k_2 \text{ of } X_2, \dots, k_n \text{ of } X_n \end{array} \right\}$$

E.g. $X, Y \sim \text{Ber}(p)$,

(1) $X = Y$

(2) X, Y independent

$$P(X=k_1, Y=k_2)$$

$$= P(X=k_1)P(Y=k_2)$$

		X	
		0	1
Y	0	$1-p$	$(1-p)^2$
	1	$(1-p)p$	$p(1-p)$
		p	p^2

$$P_{X,Y}(k_1, k_2) = P(X=k_1, Y=k_2)$$

Recovering X_j from $\underline{X} = (X_1, X_2, \dots, X_n)$: Marginals

Suppose we know $P_{\underline{X}}(\underline{k})$ for all $\underline{k} = (k_1, k_2, \dots, k_n)$.

How can we find $P_{X_1}(t)$?

E.g. Toss a fair coin twice. $X_1, X_2 \in \{0, 1\}$

$$\underbrace{\frac{1}{4} + \frac{1}{4}}_{\sim} = \frac{1}{2}.$$

$$P(X_1=0) = P(X_1=0, X_2=0) + P(X_1=0, X_2=1)$$

$$\{X_1=0\} = \{X_1=0, X_2=0\} \cup \{X_1=0, X_2=1\}$$

In general,

$$P(X_1=t) = \sum_{k_2, k_3, \dots, k_n} P(X_1=t, X_2=k_2, \dots, X_n=k_n) = \sum_{k_2, \dots, k_n} P_{\underline{X}}(t, k_2, k_3, \dots, k_n)$$

$$\{X_1=t\} = \bigcup_{k_2, \dots, k_n} \{X_1=t, X_2=k_2, X_3=k_3, \dots, X_n=k_n\}$$

$$P(X=t) = \sum_k P(x, y)(t, k)$$

(x, y)

$$P(Y=t) = \sum_k P(x, y)(k, t)$$

E.g. Toss a fair coin 3 times. $X = \# \text{ tails in first toss (0 or 1)}$
 $Y = \text{total } \# \text{ tails in all 3 (0, 1, 2, 3)}$

	Y	Ber($\frac{1}{2}$)	Outcome	X	Y
X	0 1 2 3	\downarrow $Ber(\frac{1}{2})$			
0	$\frac{1}{8}$ $\frac{1}{4}$ $\frac{1}{8}$ 0	$\rightarrow \frac{1}{2}$	HHH	0	0
1	0 $\frac{1}{8}$ $\frac{1}{4}$ $\frac{1}{8}$	$\rightarrow \frac{1}{2}$	HTT	0	1
	$\downarrow \frac{1}{8}$ $\downarrow \frac{3}{8}$ $\downarrow \frac{3}{8}$ $\downarrow \frac{1}{8}$	$\leftarrow Ber(3, \frac{1}{2})$	HTH	0	1
			HTT	0	2
			THH	1	1
			THT	1	2
			TTH	1	2
			TTT	1	3

each has
 $P = \frac{1}{8}$.

$$P(X=k, Y=l)$$

- 0 0
- 1 1
- 2 2
- 3 3

Question: Are X, Y independent?

$$P(X=1, Y=0) = 0$$

No!

$$P(X=1)P(Y=0) = \frac{1}{2} \cdot \frac{1}{8} = \frac{1}{16} \neq 0.$$

Joint distributions are just distributions of random vectors.

$$X_1, X_2, \dots, X_n \rightsquigarrow \underline{X} = (X_1, X_2, \dots, X_n)$$

possible values for \underline{X} are vectors (k_1, \dots, k_n) .

E.g. Multinomial Distribution.

Often, trials have more than 2 outcomes.

Consider a trial w r possible outcomes, w probabilities

$$p_1, p_2, \dots, p_r$$

$$p_1 + p_2 + \dots + p_r = 1$$

Perform n trials.

For $1 \leq j \leq r$, $X_j = \# \text{times we get outcome } j$.

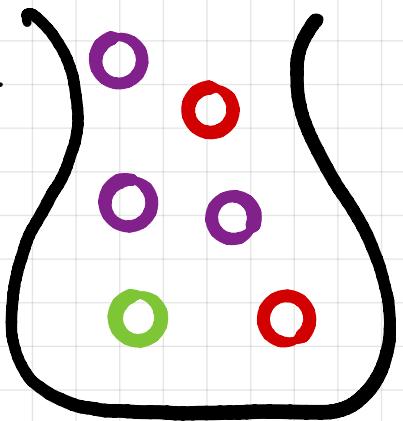
Possible values for $\underline{X} = (X_1, \dots, X_r)$: $(k_1, \dots, k_r) \leftarrow \begin{array}{l} k_j \in \{0, 1, \dots, n\} \\ k_1 + \dots + k_r = n \end{array}$

$$P(\underline{X} = \underline{k}) = \binom{\# \text{arrangements}}{w k_1, 1s, k_2, 2s, \dots, k_r, r's} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} \quad (r=2): \left(\frac{n!}{k_1! k_2!} \right) = \frac{n!}{k_1! (n-k_1)!} = \binom{n}{k_1}$$

pmf of $\text{Mult}(n; p_1, \dots, p_r)$

$$\binom{n}{k_1, k_2, \dots, k_r} = \frac{n!}{k_1! k_2! \dots k_r!}$$

E.g.



Sample 10 times with replacement.

$P(3 \text{ green}, 2 \text{ red}, 5 \text{ blue})$

$G = \# \text{green}, R = \# \text{red}, B = \# \text{blue}$

$\underline{X} = (G, R, B) \sim \text{Mult}(10; \frac{1}{6}, \frac{1}{3}, \frac{1}{2})$

$$P(\underline{X} = (3, 2, 5)) = \frac{10!}{3!2!5!} \left(\frac{1}{6}\right)^3 \left(\frac{1}{3}\right)^2 \left(\frac{1}{2}\right)^5 = 4.05\%$$

Note: if $r=2$, $\text{Mult}(n; p, q)$

E.g. Suppose $\underline{X} \sim \text{Mult}(n; p_1, p_2, \dots, p_r)$. Find the marginal distribution of X_1 .

$$P(X_1 = t) = \sum_{\substack{k_1, k_2, \dots, k_r \geq 0 \\ t + k_2 + \dots + k_r = n}} \frac{n!}{t! k_2! \dots k_r!} p_1^t p_2^{k_2} \dots p_r^{k_r}$$

$$\nearrow$$

$$k_2, k_3, \dots, k_r \geq 0$$

$$t + k_2 + \dots + k_r = n$$

Complicated! Instead observe:

$\curvearrowright X_1 = \# \text{successes in } n \text{ trials}$
 where success = outcome 1 ($P = p_1$)
 failure = outcomes $2 \dots r$ ($P = p_2 + \dots + p_r = 1 - P$)

$$X_1 \sim \text{Bin}(n, p)$$

Expectations

Let $\underline{X} = (X_1, \dots, X_n)$ be a (discrete) random vector with joint probability mass function $P_{\underline{X}}$.

If $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is a function, $Y = g(\underline{X})$ is a random variable (still discrete).

$$\begin{aligned} E(Y) &= \sum_t t \cdot P(Y=t) \\ &= \sum_k g(k) P(\underline{X}=k) \\ &= \sum_k g(k) p_{\underline{X}}(k) \end{aligned}$$

E.g. Toss a fair coin twice, $X_1, X_2 \in \{0, 1\}$.

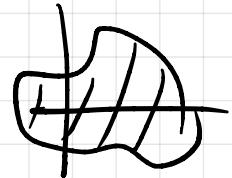
$$E(X_1 X_2) = \sum_{k_1, k_2} g(k_1, k_2) \underbrace{P(X_1=k_1, X_2=k_2)}_{1/4}$$

$$\left. \begin{aligned} Y &= 0 \cdot 0 \cancel{\frac{1}{4}} + 0 \cdot 1 \cancel{\frac{1}{4}} + 1 \cdot \cancel{0 \cdot \frac{1}{4}} \\ &\quad + 1 \cdot 1 \cancel{\frac{1}{4}} \\ &= 1/4. \approx E(X)E(Y) \end{aligned} \right\}$$

$$\left. \begin{aligned} E(X) &= 1/2 \\ E(Y) &= 1/2 \end{aligned} \right\}$$

Jointly Continuous Random Vectors

A random vector $\underline{X} = (X_1, \dots, X_n)$ has a pdf $f_{\underline{X}}: \mathbb{R}^n \rightarrow \mathbb{R}_+$ if, for "nice" subsets $B \subseteq \mathbb{R}^n$



$$P(\underline{X} \in B) = \int_B f_{\underline{X}}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

(We say X_1, X_2, \dots, X_n are jointly continuous.)

Properties: (1) $f_{\underline{X}} \geq 0$

$$(2) \int_{\mathbb{R}^n} f_{\underline{X}} = 1.$$

E.g. Standard Multivariate Normal

$$f(x_1, \dots, x_n) = (2\pi)^{-\frac{n}{2}} e^{-\frac{(x_1^2 + \dots + x_n^2)}{2}}$$

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = (2\pi)^{-\frac{n}{2}} \int_{-\infty}^{\infty} e^{-x_1^2/2} dx_1 \cdot \int_{-\infty}^{\infty} e^{-x_2^2/2} dx_2 \cdots \int_{-\infty}^{\infty} e^{-x_n^2/2} dx_n = (2\pi)^{-\frac{n}{2}} (\sqrt{\pi})^{n/2} = (2\pi)^{-\frac{n}{2}} (\sqrt{2\pi}) (\sqrt{2\pi}) \cdots (\sqrt{2\pi})$$