

LECTURE NOTES

STAT 240: PROBABILITY (ADVANCED LEVEL)

Prepared by: CALVIN KENT
www.student.math.uwaterloo.ca/~c2kent/

Instructor: CHANGBAO WU
Term: FALL 2017

Last revised: August 11, 2018

Table of Contents

	Page
PREFACE AND ACKNOWLEDGMENTS	i
CHAPTER 1 - WHAT IS PROBABILITY?	1
Random Trials	1
Random Events	1
Definition of Probability	2
Stability of Relative Frequency	4
CHAPTER 2 - PROBABILITY MODELS	6
Random Events and Sample Space	6
Discrete Probability Models	8
Classic Probability Models	10
CHAPTER 3 - CALCULATION OF CLASSIC PROBABILITIES	12
Counting Techniques	12
Probability Calculation Examples	14
CHAPTER 4 - PROBABILITY RULES	21
Venn Diagram	22
Three Axioms of Probability	25
Test 1 - Practice Sheet	27
Independent Events	30
Conditional Probability and Independence of Events	32
Conditional Probability Examples	32
Pairwise Independence	35
Multiplication Formula	39
Total (whole) Probability Formula	39
Probability Formula Examples	39
CHAPTER 5 - DISCRETE RANDOM VARIABLES AND PROBABILITY MODELS	43
Random Variables	44
Probability Mass Function (pmf)	44
Bernoulli Trials	47

Binomial Distribution	48
Geometric Distribution	49
Memoryless Property	51
Negative Binomial Distribution	52
Hyper Geometric Distribution	54
Poisson Distribution	56
Probability Distribution Examples	56
CHAPTER 6 - EXPECTATION AND VARIANCE	65
Expectation	65
Major Steps in Applications of Expectations	68
Expectation of Commonly Used Random Variables	68
General Formula on Expectations	71
Variance	71
Computation of Variance	72
Variance of Commonly Used Random Variables	74
Some Useful Results on Variance	76
Examples on Variance	76
CHAPTER 7 - DISCRETE MULTIVARIABLE DISTRIBUTIONS	78
Bivariate Distributions	78
Joint Probability Function	80
Marginal Distributions	80
Binomial Distribution (review)	83
Trinomial Distribution	83
Test 2 - Practice Sheet	87
Multinomial Distribution	90
Expectations Involving Two or More Random Variables	92
Law of Unconscious Statistician	92
Properties of Expectation	93
Covariance	94
Correlation Coefficient	95
Positive Correlation	97
Expectation and Variance of a Linear Combination of Random Variables	99
Indicator Variable Techniques	101
CHAPTER 8 - CONTINUOUS PROBABILITY DISTRIBUTIONS	104
Probability Density Function (pdf)	104
Cumulative Distribution Function (cdf)	106
Change of Variables	109
Exponential Distribution	110
Gamma Distribution	112
Uniform Distribution	114
Beta Distribution	115

Normal Distribution	116
Discretization	119
Discretization	119
Distribution Examples	120
Moments	124
Moment Generating Function (mgf)	124
Properties of MGF	127
Examples on MGF	127
Central Limit Theorem (clt)	128

Preface and Acknowledgments

This PDF document includes lecture notes for STAT 240 - Probability (Advanced Level) taught by Changbao Wu in Fall 2017.

For any questions e-mail me at `c2kent(at)uwaterloo(dot)ca`.

Be aware of the misspelling of "variance". In some parts of these notes, it was spelled as "varience" which is wrong.

Calvin KENT

Sept 08

Friday

STAT 240

1. What is probability?

Random Trials:

- Controlled experiments
 - ↳ Flip a coin
 - ↳ Card game
 - ↳ Roll a die
- Observational Studies
 - ↳ Number of traffic accidents on a certain day in Waterloo.

Outcome cannot
be predicted

- #### 2 common features for random trials:
1. The outcome cannot be predicted exactly.
 2. All the possible outcomes are known with certainty.

Random events:

Possible outcomes of a random trial.

~~Ex~~

Randomly select 2 cards from 52 card deck.

A = "Both are hearts"

B = "One heart, one spade"

C = "Both are Kings"

Event A seems to be more likely to occur than C.

- First definition of probability:

Probability is a quantitative measure of how likely a random event will occur.

→ If the trial is only run once, it is either A or not A.

→ If more trials are done (total of n times) A occurs more often than C.

- The Frequentist Definition of Probability:

- Relative frequency:

→ Trial: A random event, A

→ Repeat the trial n times, independently under the same set of conditions.

optimum number of shuffling cards (7)

Event A occurred m times.

i) $0 \leq m \leq n$

ii) $f_n(A) = \frac{m}{n} \in [0, 1]$

iii) The value of $f_n(A)$ reflects the "likelihood" of event A.

The value of $f_n(A)$ depends on the particular set of n trials.

(iv) The stability of relative frequency ($f_n(A)$)

$$n=3: m=0, 1, 2, 3$$

$$n=100: m=0, 1, \dots, 48, 49, 50, 51, 52, \dots, 99, 100$$

$$f_n(A) = \frac{m}{n} \sim P$$

Sept 11, 2017

Monday

Stability of relative frequency property

When n is large $f_n(A) = \frac{m}{n}$ tends to be very close to a fixed number, p .

• Second (frequentist) definition.

The probability of a random event A denoted $\overset{\text{as}}{=} P(A)$ is the fixed value P where $f_n(A)$ is stabilized.

Notes:

(1) " $\lim_{n \rightarrow \infty} f_n(A) = p$ " \Leftrightarrow " $f_n(A)$ stabilized at p "

$\forall \epsilon > 0, \exists N$, when $n > N$, $|f_n(A) - p| \leq \epsilon$

$$p(A) = \frac{1}{2} \quad f_n(A) = \frac{m}{n}$$

flipping a coin

(2) How to compute $P(A)$ with the frequentist definition?

not practical

(3) Frequentist interpretation of probability.

~~Ex~~

Flip a coin, $A = \text{"Head"}$ $P(A) = 0.5$

$$\frac{m}{n} \xrightarrow{\text{observed } A \text{ outcomes}} \overline{\underline{\underline{0.5}}} \quad m \text{ observed } A \text{ outcomes}$$

\hookrightarrow number of trials

~~Ex~~

Flip a coin 10 times, $A = \text{"All 10 trials are head"}$

$$P(A) \approx 0.001 = \frac{1}{2^{10}}$$

~~Ex~~

$A = \text{"Making a successful free throw"}$

$$P(A) = 0.92$$

\hookrightarrow From previous data

~~Ex~~

$B = \text{"A stat 240 student being a } \overset{\text{CS student}}{\cancel{\text{begin}}} \text{ "}$

$$P(B) = 0.95 \quad \text{different conditions}$$

~~Ex~~

$C = \text{"A new medicine is effective"}$

$$P(C) = 0.99$$

2. Discrete Probability Models

(2.1) Random events and sample space

→ Random trial

→ Simple events: Basic outcomes of the trial

→ Sample space: The set of all basic (simple) events.

→ Compound events: Consist of more than one simple event.

~~Ex~~ Flip a coin three times.

Simple events: $(HHT), (TTH), \dots$

Sample space: $\{(HHT), (TTH), (THH), \dots\}$

$A = \text{"Exactly one head"} = \{(HTT), (THT), (TTH)\}$

$B = \text{"At least two heads"} = \{(HHT), (HTH), (THH), (HHH)\}$

↓
compound event

~~Ex~~ Roll a die twice.

• Simple events: $(1,1), (1,2), \dots, (6,6)$

• Sample space: $\{(1,1), (1,2), \dots, (6,6)\}$

$A = \text{"The 1st roll turns up 5"}$

• $\{(5,1), (5,2), \dots, (5,6)\}$

$$B = \text{"The sum is 10"} = \{(4,6), (5,5), (6,4)\}$$

- Notes:
- 1) All events, simple or compound, are subsets of the sample space, S.
 - 2) Two simple events will never occur at the same time.
 - 3) A compound event occurs iff one of the "points" occurs.
 - 4) Two compound events can occur at the same time.

Sep 13, 2017

Wednesday

2.2 Discrete Probability Models

- a) The sample space, S , contains finite or countable infinite number of basic events (points)

$$S = \{a_1, a_2, \dots, a_n, \dots\}$$

- b) Let $p_i = P(a_i)$, we must have $0 \leq p_i \leq 1 \quad \forall i \sum_{i=1}^{\infty} = 1$

- c) For any event $A = \{a_{i_1}, a_{i_2}, \dots\}$

$$P(A) = p_{i_1} + p_{i_2} + \dots$$

~~Ex~~
Roll a die.

$$S = \{1, 2, 3, 4, 5, 6\}$$

a_i	1	2	3	4	5	6
p_i	p_1	p_2	p_3	p_4	p_5	p_6

~~Ex~~
Roll a die n times (independently)

Let $m_i = \#$ times "i" turns up where $i \in \{1, 2, 3, 4, 5, 6\}$

$$m_1 + m_2 + \dots + m_6 = n \Rightarrow \frac{m_1}{n} + \frac{m_2}{n} + \dots + \frac{m_6}{n} = 1$$

$$\downarrow \quad \downarrow \quad \downarrow$$

$$p_1 \quad p_2 \quad \dots \quad p_6$$

~~Bx~~
 $B = \text{"An even # turns up"}$
 $= \{2, 4, 6\}$

$$P(B) = ?$$

$$f_n(B) = \frac{M_2 + M_4 + M_6}{n}$$

relative frequency
in n trials

Probability versus Statistics:

- Probability starts with an assumed model, and it studies properties of the model.
- Statistics starts with an observed data (sample) tries to find a suitable model for the problem.
- All models are wrong but some are useful.

- George Box

~~Bx~~
 $S = \{0, 1, 2, \dots\}$

$$p_i = P(\{i\}) = 0.6^i \cdot c, i = 0, 1, 2, \dots$$

a) $c = ?$

~~q = 0.6~~ $\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$ $q = 0.6$ then

$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-0.6} = \frac{10}{4}$$

$$c \cdot \frac{10}{4} = 1 \text{ then } \underline{\underline{c = 2.5^{-1}}}$$

b) $B = \text{"The outcome number } \leq 10\text{"}$

$$P(B) = \sum_{i=0}^{10} p_i = \sum_{i=0}^{10} c \cdot 0.6^i$$

$$\sum_{i=0}^{10} q^i = \frac{1-q^{10+1}}{1-q} \quad \therefore \sum_{i=0}^m q^i = \frac{1-q^{m+1}}{1-q}$$

c) $C = \text{"The # is odd"}$

$$C = \{1, 3, 5, 7, \dots\} \quad P(C) = \sum_{j=1}^{\infty} c \cdot 0.6^{2j-1}$$

$$= \sum_{j=1}^{\infty} \frac{c}{0.6} (0.36)^j = ?$$

2.3 Classic Probability Models

(i) S has a finite number of points

$$S = \{a_1, a_2, \dots, a_N\}$$

(ii) All basic events are equally likely to occur.

$$\text{i.e. } P(a_1) = P(a_2) = \dots = P(a_N) = \frac{1}{N}$$

(iii) $A = \{a_{i_1}, a_{i_2}, \dots, a_{i_M}\} \Rightarrow P(A) = \frac{M}{N}$

$$= \frac{\text{number of points in } A}{\text{number of points in } S}$$

~~Ex~~ Flip a coin 10 times, independently.

(i) $S = \{(HHTHT\ldots T), (THT\ldots T), \dots\}$ $N = 2^{10}$

(ii) $A = \text{"All 10 are Heads"}$

$$P(A) = \frac{1}{2^{10}} \approx 0.001$$

(iii) $C = \text{"There are exactly 3 Heads among the 10"}$

$$P(C) = \frac{M}{N} = \frac{\text{Number of favorable outcomes}}{2^{10}}$$

Then it should be 3H and 7T

Sep 15, 2017
Friday

3- Calculation of Classic Probabilities

3.1. Counting Techniques

1) Full factorial: The number of ways to put n distinguishable objects in a row.

$$n! = n(n-1)(n-2)\dots(2)(1)$$

2) Permutations: The number of ways to ~~put~~ select r objects from n distinguishable objects and put them in a row.

$$n=4, r=2$$

$$P_n^r = n^{(r)} = \frac{n(n-1)\dots(n-r+1)(n-r)}{(n-r)!}$$

$$= n(n-1)\dots(n-r+1)$$

3) Combinations: The number of ways to select r objects from n distinguishable objects and put them into a container.

↳ Order does not matter.

↳ Grab a handful of r .

$$C_n^r = \binom{n}{r} = \frac{n^{(r)}}{r!} = \frac{n!}{(n-r)!r!}$$

$$\binom{n}{r} = \binom{n}{n-r}$$

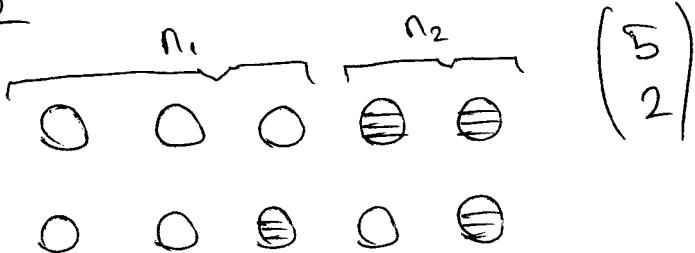
~~$$\binom{15}{3} = \frac{15 \cdot 14 \cdot 13 \cdot 12!}{12! \cdot 3!} = \frac{5 \cdot 7 \cdot 13}{\cancel{12!} \cdot \cancel{3!}} = 455$$~~

4) Seat-reserving: (Combination number)

The number of ways to arrange n objects in a row,
among them n_1 objects are identical of one type,
the other $n_2 = n - n_1$ objects are identical of another type

$$n = 5, \quad n_1 = 3, \quad n_2 = 2$$

$$\frac{n!}{n_1! n_2!} = \binom{n}{n_1}$$

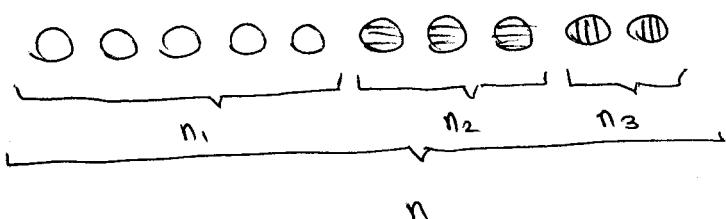


Generalization: $\forall n = n_1 + n_2 + \dots + n_k$

n_1 objects are identical of type 1;
 \vdots
 n_k objects " " " " " = k ;

How many ways we can arrange them in a row?

$$n = 10; \quad n_1 = 5, \quad n_2 = 3, \quad n_3 = 2$$



$$\frac{10!}{5! 3! 2!} = \frac{n!}{n_1! n_2! n_3!}$$

$\therefore n_1$ (and so os n_2 and n_3) can be arranged within their own groups. Since they're identical (among their own groups) order doesn't matter

$$\frac{n!}{n_1!n_2!n_3!} = \binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3}$$

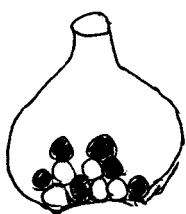
冗余的

~~B2+~~ (continued) Flip a coin 10 times; $P(A) = ?$
Where $A = \{\text{Exactly 3 Heads}\}$

$$P(A) = \frac{M}{N} = \frac{\binom{10}{3}}{2^{10}}$$

(Reserve "3 Heads" among
10 flips. Since HHH TTT TTTT
is equiv. to HTHHTTTTTT
use combination.)

~~B2+~~ An urn contains 6 black, 5 white balls. All
distinguishable. Randomly select 3 without replacement.



- A: No blacks
- B: Exactly two blacks
- C: At least two blacks
- D: The 2nd draw is black

Solution 1: Take 3 with big hand.

$$P(A) = \frac{\binom{5}{3}}{\binom{11}{3}}$$

explained:

$$P(B) = \frac{\binom{6}{2} \cdot \binom{5}{1}}{\binom{11}{3}}$$

and

$\binom{5}{3} \rightarrow$ Take 3 from all whites

Sept 18, 2017
Monday

- ① 6 Black, 5 White. Randomly select 3 without replacement.

A = "No black"

B = "Exactly two black"

C = "At least two black"

D = "The 2nd draw is black"

$$\cdot P(A) = \frac{\binom{5}{3}}{\binom{11}{3}}$$

↙ all possibilities

Big hand of 3.

$$\cdot P(B) = \cancel{\frac{\binom{6}{2} \binom{5}{1}}{\binom{11}{3}}}$$

$$\cdot P(C) = \frac{\binom{6}{2} \binom{5}{1} + \binom{6}{3}}{\binom{11}{3}}$$

or 1 - $\frac{\binom{5}{3} + \binom{5}{2} \binom{6}{1}}{\binom{11}{3}}$

$$\cdot P(A) = \frac{\binom{3}{3}}{\binom{11}{3}}$$

seat reserving

$$\cdot P(B) = \frac{\binom{3}{2} 5^{11} 6^{(2)}}{\binom{11}{3}}$$

$$\cdot P(C) = \frac{\cancel{\binom{3}{2}} \cancel{\binom{3}{1}} 5^{11} 6^{(2)} + \binom{8}{3} 6^{(3)}}{\binom{11}{3}}$$

Order is considered

$$P(D) = \frac{\binom{6}{1} 10^{(2)}}{\binom{11}{3}}$$

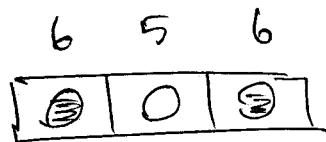
①

* $n=3, r=2$ then ${}^3P_2 = 6$ and $\binom{3}{2} = 3$

2) 6 Black, 5 White. Pick 3 with replacement

$$P(A) = \frac{5^3}{11^3}$$

$$P(B) = \frac{5 \cdot 6^2 \cdot 3}{11^3} \quad \text{↑ } \binom{3}{2} \text{ seat reserving}$$



3) Suppose one's birthday is equally likely in one of the 12 months. Consider a group of 9 people.

A = "No Bday in Jan or in Dec."

B = "All at diff months"

C = "All in the same month"

D = "2 in one month, 3 in another, rest in four other diff. months"

$$\cdot P(A) = \frac{10^9}{12^9}$$

$$\cdot P(C) = \frac{12}{12^9}$$

$$\cdot P(B) = \frac{12^9}{12^9}$$

$$\cdot P(D) =$$

$$\cdot P(B) = \frac{\binom{12}{9} \cdot 9!}{12^9}$$

Req. represent representation
of at least 2 ppl having same
bday of 23 ppl group? (2)

4) Roll a die 4 times

A: "The sum is 10"

$$P(A) = \frac{\text{Number of favorable outcomes}}{6^4} \rightarrow \begin{array}{ccc} 6 & 2 & 11 \\ 5 & 3 & 11 \\ ; & & \end{array}$$

0 0 0 0 1 0 0 1 0 1 0, Sum is 10
10 balls

Distribute 10 balls
in 4 sections

Observations:

- Each section must have ~~at least~~ at least one ball.
- Arrange 3 bars to make 4 sections.

$$P(A) = \frac{\binom{9}{3} - \binom{4}{1}}{6^4} \rightarrow \begin{array}{cccc} 1 & 1 & 1 & 7 \\ 1 & 1 & 7 & 1 \\ 1 & 7 & 1 & 1 \\ 7 & 1 & 1 & 1 \end{array}$$

Die cannot have 7

or ~~permutation theorem~~ multinomial theorem

Sept 20, 2017
Wednesday

Example

Birth month; 9 students

D = "2 in one month; 3 in another month, the remaining 4 in four other different months."

E = "3 in one month; 3 in another month, 3 in a third month."

Remark: If the trial requires step 1 AND step 2 AND step 3, etc. to finish, use multiplication.

If the trial is finished under scenario (i) OR scenario (ii) but not both.

$$P(D) = \frac{\binom{12}{1} \binom{9}{2} \binom{11}{1} \binom{7}{3} \cdot 10^{(4)}}{\cancel{12^9}}$$

$\binom{12}{1}$: Pick a month

$\binom{9}{2}$: Pick 2 students

$\binom{11}{1}$: Pick another month

$\binom{7}{3}$: Pick 3 other students

$10^{(4)}$: Distribute remaining 4 students among 10 months.

$$\text{or } \frac{\binom{12}{1} \binom{9}{2} \binom{11}{1} \binom{7}{3} \cdot \binom{10}{4} \cdot 4!}{\cancel{12^9}}$$

$$P(E) = \frac{\binom{12}{3} \binom{9}{3} \binom{6}{3} \binom{3}{3}}{12^9}$$

Compare with

$$\frac{\binom{12}{1} \binom{9}{3} \binom{11}{1} \binom{6}{3} \binom{10}{1} \binom{3}{3}}{12^9}$$

Consider with 2 students.

$$\begin{array}{c} \frac{12}{3} \quad \frac{9}{3} \quad \frac{6}{3} \quad \frac{3}{3} \\ | \qquad | \qquad | \qquad | \\ 8 \text{ months} \quad 3 \text{ students in one of the 3 months} \end{array}$$

3 students in another month in 3 months

(1)

(5) Randomly select 3 numbers without replacement from $\{0, 1, 2, \dots, 9\}$ and form a 3-digit sequence.

$A =$ "The sequence is a 3-digit even number."

$$P(A) = \frac{\binom{5}{1}\binom{5}{1}\binom{8}{1} + \binom{4}{1}\binom{4}{1}\binom{8}{1}}{10^3}$$

3 digit sequences

Two scenarios:

- ① First digit is from: $\{1, 3, 5, 7, 9\}$
- ② or $\{2, 4, 6, 8\}$

$$\textcircled{1} \quad \begin{matrix} (5) & (5) & (8) \\ 1^{\text{st}} \text{ digit} & 3^{\text{rd}} \text{ digit} & 2^{\text{nd}} \text{ digit} \end{matrix} \quad + \quad \textcircled{2} \quad \begin{matrix} (4) & (4) & (8) \\ 1^{\text{st}} \text{ digit} & 3^{\text{rd}} \text{ digit} & 2^{\text{nd}} \text{ digit} \end{matrix}$$

(6) Select 3 numbers with replacement from $\{0, 1, \dots, 9\}$ and form a 3-digit sequence.

$B =$ "The three numbers are in strict increasing order."

$C =$ "Sum is 10"

$$P(B) = \frac{\binom{10}{3}}{10^3}$$

Three in a particular order
 \iff Three with no order

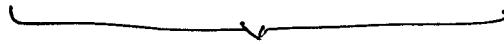
$$P(C) = \frac{\text{Number of favorable outcomes}}{10^3}$$

The sum is 10.



2 "moving bars"

0 0 0 0 0 0 0 0 0



10 numbers

$$\begin{array}{r} \binom{12}{10} - 3 \\ \hline \cancel{\#} 10^3 \end{array}$$

because number "10" cannot
be picked

Sept 22, 2017
Friday

~~Px~~
Flip a coin twice.

$$A: "1^{\text{st}} \text{ is head}" = \{(HH), (HT)\}$$

$$B: "2^{\text{nd}} \text{ is head}" = \{(TH), (HH)\}$$

$$C: "Exactly one head" = \{(HT), (TH)\}$$

$$D: "Both are heads" = \{(HH)\}$$

$$E: "At least one head" = \{(HT), (TH), (HH)\}$$

$$F: "No heads" = \{(TT)\}$$

$$G: "Three heads" = \emptyset$$

4. Probability Rules

4.1 Relations among events

① Any event is a subset of S.

S: An event that always occurs

\emptyset : An event that will never occur.

② Union of events:

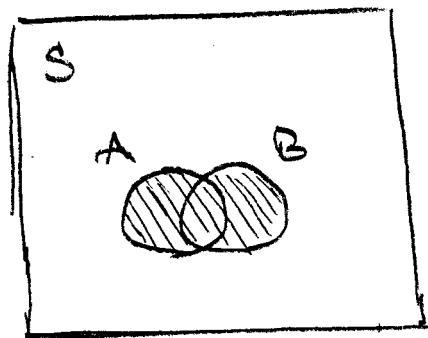
$A \cup B$: A or B (at least one of the two)

i.e. $A \cup B = E$

$C \cup D = E$

①

Venn Diagram:

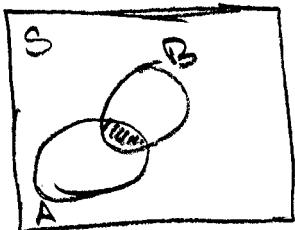


• $A \cup B \cup C$: A or B or C (at least one of the three)

$A_1 \cup A_2 \cup \dots \cup A_n$: At least one of A_i ; $1 \leq i \leq n$

③ Intersection:

$A \cap B$: A and B (Both)



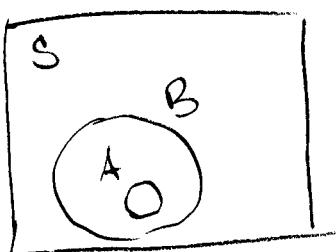
$$A \cap B = AB$$

④ Mutually exclusive (disjoint)

A and B are mutually exclusive if A and B cannot happen at the same time.

$$A \cap B = \emptyset$$

⑤ Inclusion: $A \subset B$: A is included in B
 $B \supset A$: B includes A

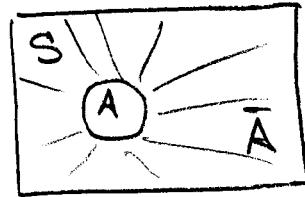


B occurs whenever A occurs.

2

⑥ Opposite event (complement)

\bar{A} : not \bar{A} ; $\bar{A} = S \setminus A$



Properties: a) $\emptyset \subset A \subset S$

b) $A \cup A = A$, $A \cap A = A$

$A \cup S = S$, $A \cap S = A$

$A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$

c) if $A \subset B$ then $A \cup B = B$
 $A \cap B = A$

d) Distribution Law: $A(B \cup C) = AB \cup AC$

Use Venn diagram to show distribution law.

- Words

e) Morgan's Laws: $\overline{A \cup B} = \overline{A} \cap \overline{B}$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$\overline{ABC} = \overline{A} \cup \overline{B} \cup \overline{C}$$

$$f) \overline{(\bar{A})} = A$$

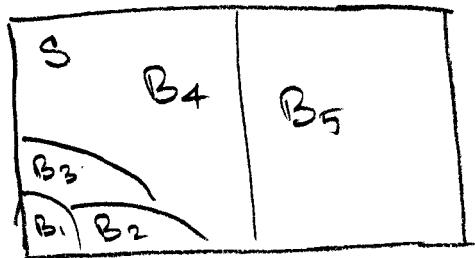
$$g) A \cup \bar{A} = S$$

a partition of S

$$\overline{A \bar{B}} = \bar{A} \cup B$$

$$A \bar{A} = \emptyset$$

h) B_1, B_2, \dots, B_n form a partition of S if



$$B_1 \cup \dots \cup B_n = S$$

$$B_i \cap B_j = \emptyset : i \neq j$$

Probability Rules:

$P(\cdot)$ is a real-valued set function defined over S , satisfying

- Three axioms of probability
- (i) Non-negativity: $P(A) \geq 0$ for any $A \subset S$
 - (ii) Normality: $P(S) = 1$
 - (iii) Additivity: For any sequence of mutually exclusive events A_1, A_2, \dots with $A_i \cap A_j = \emptyset$ for $i \neq j$

$$\text{then } P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

• Other properties of probability.

$$(1) P(\emptyset) = 0$$

$$P(\emptyset) + P(\emptyset) + \dots = P(\emptyset) \because \emptyset \cup \emptyset = \emptyset \text{ and (iii)}$$

$$(2) A_1, A_2, \dots, A_n: A_i \cap A_j = \emptyset \text{ for } i \neq j$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

from (iii) and (1)

$$3) P(\bar{A}) = 1 - P(A) \because A \bar{A} = \emptyset, A \cup \bar{A} = S \text{ and (ii) and (iii)}$$

$$4) P(B) = P(AB) + P(\bar{A}B) \because B = AB \cup \bar{A}B \text{ and } (AB) \cap (\bar{A}B) = \emptyset$$

$$5) \text{ If } A \subset B \text{ then } P(A) \leq P(B)$$

$$\therefore A \subset B \Rightarrow B = (A \cup \bar{A}B) \Rightarrow P(B) = P(A) + P(\bar{A}B)$$

(6) For any two events A and B, $P(A \cup B) = P(A) + P(B) - P(AB)$

$$A \cup B = A \cup (\bar{A}B) \Rightarrow P(A \cup B) = P(A) + P(\bar{A}B)$$

$$B = (AB) \cup (\bar{A}B) \Rightarrow P(B) = P(AB) + P(\bar{A}B)$$

↳ For any events A, B, C,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

↳ For any n number of events A_1, A_2, \dots, A_n

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i A_j) + \sum_{i < j < k} P(A_i A_j A_k) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n)$$

↳ Symmetric Case

$$P\left(\bigcup_{i=1}^n A_i\right) = n P(A_1) - \binom{n}{2} P(A_1 A_2) + \binom{n}{3} P(A_1 A_2 A_3) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n)$$

~~Bx~~
An urn contains n black balls and m white balls, all distinguishable. Randomly select r balls without replacement where $r \leq n$ and $r \leq m$.

A = "Getting at least 1 black ball."

$$A = 1 - \frac{\binom{m}{r}}{\binom{n+m}{r}} \quad \text{or} \quad A = \frac{\binom{n}{1}\binom{m}{r-1} + \binom{n}{2}\binom{m}{r-2} + \dots + \binom{n}{r}\binom{m}{0}}{\binom{n+m}{r}}$$

$$\binom{n+m}{r} = \sum_{i=0}^r \binom{n}{i} \binom{m}{r-i}$$

Classical Definition: $\frac{\# \text{ways the event can occur}}{\# \text{outcomes in total}}$

outcomes in total: $|S|$, S = sample space

Frequentist Definition: The likeliness of an event A occurring in long number of trials $\left(\frac{m}{n}\right) = f_n(A)$

- Not practical
- Requires n trials

Deterministic Model: Assumes certainty.

A given input produces same output.

Does not rely on random variables.

S: Sample Space: Set of all distinct outcomes for an experiment.

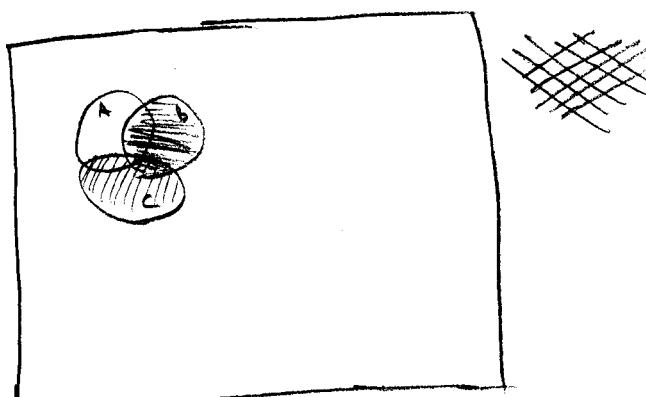
Simple event: Single element in S.

Compound event: An event in S that is made up of two or more simple events.

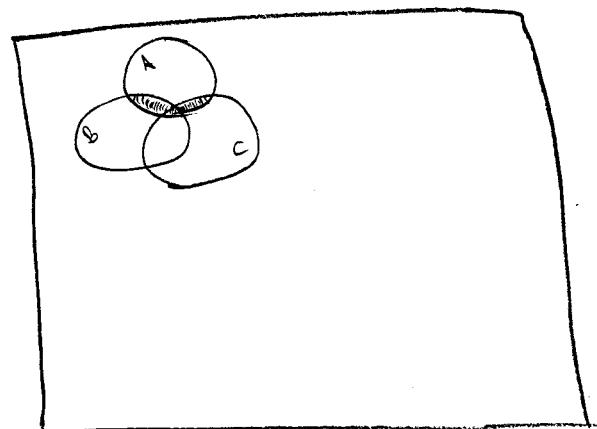
$$A(B \cup C) = AB \cup AC$$

$$A(B \cup C) = \{x : x \in A \text{ and } x \in B \cup C\}$$

WLOG suppose the Venn diagram of sets A, B and C is:



then show $A(B \cup C)$



then show $AB \cup AC$

Sep 27 2017
Wednesday

~~Bx~~ $P(A) = 0.4, P(B) = 0.7, P(A \cup B) = 0.8$

1) $P(AB) = ?$

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

then $P(AB) = -(0.8 - 0.4 \cdot 0.7) = \underline{\underline{0.3}}$

2) $P(\bar{A} \bar{B}) = 1 - P(\bar{A} \bar{B}) = 1 - P(A \cup B) = \underline{\underline{0.2}}$

$$P(\bar{A} \cup \bar{B}) = 1 - P(AB)$$

3) $P(\bar{A} \bar{B}) = 0.4$

Recall: $P(B) = P(AB) + P(\bar{A}B)$

4) $P(\bar{A} \cup B) = 0.9$

use i) $P(\bar{A} \cup B) = 1 - P(\bar{A} \bar{B})$

or ii) $P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A}B)$

~~Eg~~ Randomly take a hand of 13 cards from a deck of 52.

B: "The hand has 6 spades or one ace"

A_1 : "6 spades" A_2 : "1 ace" A_3 : "6 spades, one of six is an ace"

then $A_1 A_2 = A_3$

$$P(B) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2)$$

$$= \frac{\binom{13}{6} \binom{39}{7} + \binom{4}{1} \binom{48}{12} - \binom{1}{1} \left(\binom{12}{5} \binom{36}{7}\right) - \binom{3}{1} \binom{13}{6} \binom{36}{6}}{\binom{52}{13}}$$

~~Ex~~ Roll a die 3 times. B = "The #6 appears at least once."

i) $1 - P(B) = P(B) = 1 - \frac{5^3}{6^3} = \frac{91}{216}$

B_1 : "The #6 appears on 1st roll"

B_2 " = 2nd roll"

B_3 " = 3rd roll"

ii) $P(B) = P(B_1 \cup B_2 \cup B_3) = P(B_1) + P(B_2) + P(B_3) - P(B_1B_2) - P(B_1B_3) - P(B_2B_3) + P(B_1B_2B_3)$
 $= 3P(B_1) - 3P(B_1B_2) + P(B_1B_2B_3)$

\therefore Probabilities are symmetric.

$$= 3 \cdot \frac{1}{6} - 3 \cdot \frac{1}{36} + \frac{1}{216} = \frac{108 - 18 + 1}{216} = \underline{\underline{\frac{91}{216}}}$$

iii) C_i : "The #6 appears exactly i times."

$$B = C_1 + C_2 + C_3 \Rightarrow P(B) = P(C_1 \cup C_2 \cup C_3)$$

$$C_2: \frac{\binom{1}{1}\binom{1}{1}\binom{5}{1} \cdot \binom{3}{2}}{6^3}$$

~~Ex~~ The matching problem. n letters: $1, 2, 3, \dots, n$
 n envelopes: $1, 2, 3, \dots, n$

Randomly put the n letters into the n envelopes, one letter per envelope.

$$\frac{P(AB)}{P(B)} = P(A) \Rightarrow A \text{ and } B \text{ are independent}$$

Defn: Two events A and B are independent if

$$P(AB) = P(A) \cdot P(B)$$

Notes: (i) A and B are ind. $\Rightarrow P(A|B) = P(A)$
and
 $P(B|A) = P(B)$

(ii) Relation to "A and B mutually exclusive":

$AB = \emptyset \Rightarrow A \text{ and } B \text{ are mutually exclusive}$

If $P(A) \neq 0, P(B) \neq 0$ and A and B are mutually exclusive, then A and B are NOT ind.

$P(A) \cdot P(B) \neq 0$ then $P(AB) \neq 0$ then $P(A) \cdot P(B) \neq P(AB)$

(iii) If A and B are ind. then

A and \bar{B} are ind.

\bar{A} and B are ind.

A and \bar{B} are ind.

If $P(AB) = P(A)P(B)$ then $P(A\bar{B}) = P(A) \cdot P(\bar{B})$

$$\begin{aligned} P(A\bar{B}) &= P(A) - P(AB) = P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \end{aligned}$$

$$P(A\bar{B}) = P(A)P(\bar{B})$$

~~A and B are ind.~~

$$P(AB) = P(A)P(B), P(A\bar{B}) = P(A)P(\bar{B}), P(\bar{A}B) = P(\bar{A})P(B)$$

$$P(\bar{A}\bar{B}) = P(\bar{A})P(\bar{B})$$

~~(*) A and B are ind. $P(A) = 0.2, P(B) = 0.4$~~

a) $P(A \cup B) = 1 - P(\bar{A}\bar{B}) = 1 - P(\bar{A})P(\bar{B}) = 1 - 0.8 \cdot 0.6 = 0.52$

$$P(A \cup B) = P(A) + P(B) - P(A)P(B) = 0.6 - 0.08 = 0.52$$

b) $P(A \cup \bar{B}) = 1 - P(\bar{A}B) = 1 - 0.8 \cdot 0.4 = 0.68$

(iv) Independence comes from independent random trials.

~~(*) Two kids write an exam independently.~~

A_1 : "Kid 1 passes the exam"

A_2 : "Kid 2 passes the exam"

$P(A_1) = 0.8, P(A_2) = 0.9, A_1$ and A_2 are ind.

B: "Both pass the exam"

C: "At least one passes the exam"

D: "Exactly one passes the exam"

$$B = A_1 A_2 = 0.8 \cdot 0.9$$

$$D = C - B$$

$$C = A_1 \cup A_2$$

$$D = A_1 \bar{A}_2 \cup \bar{A}_1 A_2$$

Sept 29, 2017
Friday

4.3. Conditional Probability and Independence of Events

(1) Conditional Probability: The probability of A given that B has already occurred.

The conditional probability of A given B;

$$P(A|B) = \frac{P(AB)}{P(B)}, P(B) \neq 0$$

$$n=6: \quad \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \quad f_n(A) = \frac{3}{6}$$

$$\begin{matrix} A & \bar{A} & A & A & \bar{A} & \bar{A} \end{matrix}$$

$$\begin{matrix} \bar{B} & B & B & B & \bar{B} & B \end{matrix}$$

$$f_n(B) = \frac{4}{6}$$

$$f_n(A|B) = \frac{2}{4} = \frac{\frac{2}{6}}{\frac{4}{6}} = \frac{f_n(AB)}{f_n(B)} f_n(AB) \cdot \frac{2}{6}$$

~~Ex~~
 $P(A) = 0.4, P(B) = 0.7, P(A \cup B) = 0.8$

a) $P(A|B) = \frac{P(AB)}{P(B)}$, b) $P(A|\bar{B}) = \frac{P(A\bar{B})}{P(\bar{B})}$ c) $P(\bar{A}|\bar{B}) = \frac{P(\bar{A}\bar{B})}{P(\bar{B})}$

Remark: $P(\bar{A}) = 1 - P(A)$ and $P(\bar{A}|B) = 1 - P(A|B)$

• Show $P(\bar{A}|B) = 1 - P(A|B)$:

$$\text{L.H.S: } \frac{P(\bar{A}B)}{P(B)} \quad \text{and R.H.S: } 1 - \frac{P(AB)}{P(B)}$$

~~Ex~~ Randomly take a hand of 13 cards from a deck of 52.

A: "The hand has exactly 7 spades"

B: "The hand has all 4 aces"

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{P(AB)}{\binom{(4)}{4} \binom{48}{9}}}{\frac{\binom{(4)}{4} \binom{48}{9}}{\binom{52}{13}}} = \frac{\binom{4}{4} \binom{12}{6} \binom{36}{3}}{\binom{4}{4} \binom{48}{9}}$$

(2) Two events A and B are independent

~~Ex~~ Roll a die twice.

A = "1st roll is 5", B = "2nd roll is 4", C = "Sum is 10"

$$P(A|B) = P(A) = \frac{1}{6}$$

$$P(C|A) \neq P(C)$$

A = "At least one letter matches its envelope"

B_i = "Letter i matches envelope i" where i = 1, 2, 3, ..., n

P(A) = P(B₁ ∪ B₂ ∪ ... ∪ B_n) where B₁, B₂, ..., B_n are symmetric

$$= (-1)^0 \binom{n}{1} P(B_1) + (-1)^1 \binom{n}{2} P(B_1 B_2) + (-1)^2 \binom{n}{3} P(B_1 B_2 B_3) + \dots + (-1)^{n-1} \binom{n}{n} P(B_1 B_2 \dots B_n)$$

$$= (-1)^0 \frac{1}{n!} + (-1)^1 \binom{n}{2} \frac{1}{n(n-1)} + (-1)^2 \binom{n}{3} \frac{1}{n(n-1)(n-2)} + \dots + (-1)^{n-1} \frac{1}{n!}$$

$$= (-1)^0 \cdot 1 + (-1)^1 \frac{1}{2!} + (-1)^2 \frac{1}{3!} + \dots + (-1)^{n-1} \frac{1}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots + (-1)^{n-1} \frac{1}{n!}$$

$$P(\bar{A}) = P(\text{no match}) = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \dots + (-1)^n \frac{1}{n!}$$

C = "Only the first r letters match their envelopes."

D = "Exactly r letters match their envelopes."

Oct 2, 2017

Monday

(3) Three or more events independent of each other

A, B, C are independent if

$$P(AB) = P(A)P(B) \quad (\text{i})$$

$$P(AC) = P(A)P(C) \quad (\text{ii})$$

$$P(BC) = P(B)P(C) \quad (\text{iii})$$

$$P(ABC) = P(A)P(B)P(C) \quad (\text{iv})$$

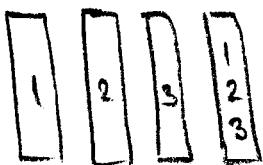
Monte Hall problem

(i)(ii)(iii) : pairwise independence

(i), (ii), (iii) \Rightarrow (iv)

~~Example~~

4 cards



Randomly take a card.
 A_i = "The card has # i on it"

$$P(A_1, A_2) = P(A_1) \cdot P(A_2)$$

$$\frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2}$$

but

$$P(A_1, A_2, A_3) \neq P(A_1) P(A_2) P(A_3)$$

$$\frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$

- n events A_1, A_2, \dots, A_n indep:

$$P(A_i; A_j) = P(A_i)P(A_j), i \neq j$$

$$P(A_i; A_j; A_k) = P(A_i)P(A_j)P(A_k), i < j < k$$

⋮
⋮

$$P(A_1; A_2; \dots; A_n) = P(A_1)P(A_2)P(A_3) \dots P(A_n)$$

similarly $P(\bar{A}_i; A_j) = P(\bar{A}_i)P(A_j)$

$$P(A_i; \bar{A}_j; \bar{A}_k) = P(A_i)P(\bar{A}_j)P(\bar{A}_k)$$

~~Example~~

The prob. of Raptors winning a game against Cavs is 0.40.
Games are played independently.

A_i : "Raptors win game i" $i=1, 2, 3, \dots$

$$P(A_i) = 0.40$$

A_1, A_2, \dots, A_n are independent of each other.

a) B : "Raptors win at least one game among three games played"

$$B = A_1 \cup A_2 \cup A_3 \quad \text{or} \quad B = \underbrace{A_1 \bar{A}_2 \bar{A}_3 \cup \bar{A}_1 A_2 \bar{A}_3 \cup \bar{A}_1 \bar{A}_2 A_3}_{\text{exactly one}} \cup \underbrace{\bar{A}_1 \bar{A}_2 \bar{A}_3}_{\text{exactly two}} \cup \dots$$

each event is
independent

$$P(A_1 \bar{A}_2 \bar{A}_3) = 0.4 \cdot 0.6^2 + P(\bar{A}_1 A_2 \bar{A}_3) = 0.6 \cdot 0.4 \cdot 0.6 + \dots + P(A_1 A_2 A_3) = 0.4^3$$

$$\text{or } P(A_1 \cup A_2 \cup A_3) = 1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = 1 - \underline{\underline{0.6^3}}$$

b) C = "Raptors win a best of seven series in 5 games."

$$C = \bar{A}_1 A_2 A_3 A_4 A_5 \cup A_1 \bar{A}_2 A_3 A_4 A_5 \cup A_1 A_2 \bar{A}_3 A_4 A_5 \cup A_1 A_2 A_3 \bar{A}_4 A_5$$

$$P(C) = 0.6 (0.4)^4 \cdot 4$$

$$= \left[\binom{4}{1} 0.6 \cdot 0.4^3 \right] \cdot 0.4$$

c) The prob. that Raptors win a best of seven series.

$$\binom{4}{0} 0.4^4 + \binom{4}{1} \cdot 0.4^3 \cdot 0.6 \cdot 0.4 + \binom{5}{2} \cdot 0.4^3 \cdot 0.6^2 \cdot 0.4 + \binom{6}{3} 0.4^3 \cdot 0.6^3 \cdot 0.4$$

↓ ↓ ↙ ↙ ↙ ↙ ↓
 lose 0 lose 1 5th game is won 6th game is won 7th game is won
 ↙ ↙ ↙ ↙ ↙
 lose 2 lose 3

d) Games are played until one team wins two games in a row which wins the series.

D: "Raptors win the series"

$$D = A_1 A_2 \cup A_1 \bar{A}_2 A_3 A_4 \cup A_1 \bar{A}_2 A_3 \bar{A}_4 A_5 A_6 \cup \dots \cup \bar{A}_1 A_2 A_3 \cup \bar{A}_1 A_2 \bar{A}_3 A_4 A_5 \cup \dots$$

All mutually exclusive. Then

$$P(D) = 0.4^2 + 0.6 \cdot 0.4^3 + 0.6^2 \cdot 0.4^4 + \dots + 0.6 \cdot 0.4^2 + 0.6^2 \cdot 0.4^3 + 0.6^3 \cdot 0.4^4$$

$$0.4^2 \left(\sum_{i=0}^{\infty} 0.4^i 0.6^i \right) + 0.6 \cdot 0.4^2 \left(\sum_{i=0}^{\infty} 0.4^i 0.6^i \right)$$

check?

Two

4.4) ~~Three~~ formulas for probabilities:

$$\textcircled{1} \quad P(A|B) = \frac{P(AB)}{P(B)} \Rightarrow P(AB) = P(B) \cdot P(A|B)$$
$$P(AB) = P(A) \cdot P(B|A)$$

Multiplication formula

\textcircled{2} The total probability formula
(whole) Type your text

$$P(B) = P(AB) + P(\bar{A}B)$$

$$P(B) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})$$

Oct 4, 2017

Wednesday

(1) The multiplication (product) formula

$$P(AB) = P(A)P(B|A)$$

$$P(A_1 A_2 \dots A_n) = P(A_1) \underbrace{P(A_2 | A_1)}_{\frac{P(A_1 A_2)}{P(A_1)}} \underbrace{P(A_3 | A_1 A_2)}_{\frac{P(A_1 A_2 A_3)}{P(A_1 A_2)}} \dots P(A_n | A_1 A_2 \dots A_{n-1})$$

~~Ex~~

An urn contains 3 black and 4 white

Randomly select one ball. The ball along with another

ball of the same colour is put back; restart the trial again.

B: "The first three selections are all black."

A_i: "The ith selection is black" i=1,2,3

$$P(B) = P(A_1 A_2 A_3)$$

$$= P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) = \frac{3}{7} \cdot \frac{4}{8} \cdot \frac{5}{9}$$

(2) The total (whole probability formula

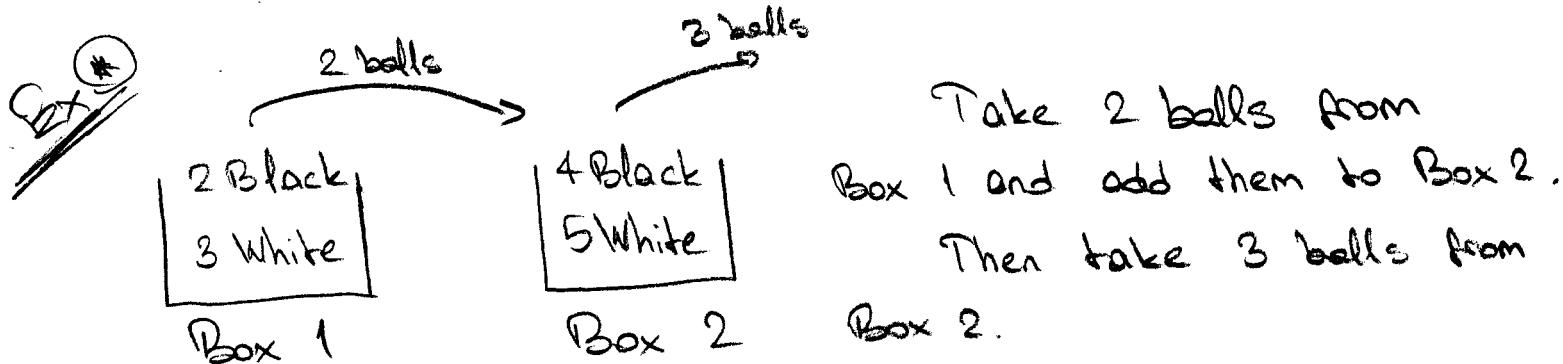
$$P(B) = P(AB) + P(\bar{A}B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$

Generalization: A₁, A₂, ... form a partition of S.

$$A_1 \cup A_2 \cup \dots = S \quad \text{and} \quad A_i \cap A_j = \emptyset, \quad i \neq j$$

$$P(B) = P(B|S) = \sum_{i=1}^{\infty} P(A_i; B) = \sum_{i=1}^{\infty} P(A_i)P(B|A_i)$$

where $P(B|A_i)$: "Forward" conditional probabilities



B: "The three balls selected from box 2 are all black."

A_i : "There are i black balls selected from Box 1" $i=0,1,2$

$$\begin{cases} A_0 \cup A_1 \cup A_2 = S \\ A_i \cap A_j = \emptyset \quad i \neq j \end{cases} \quad \rightarrow \text{forms a partition}$$

$$P(B) = \sum_{i=0}^2 P(A_i)P(B|A_i) = \sum_{i=0}^2 \frac{\binom{2}{i}}{\binom{5}{2}} \cdot \frac{\binom{4+i}{3}}{\binom{11}{3}} ; \quad \binom{4+i}{3} \text{ := Pick 3 balls from } 4+i \text{ black balls}$$

Chapter 4, question 16

Roll 2 dice.

- 1) Win immediately if total number of sides is 7 or 11
- 2) Lose immediately if " " " " " 2, 3, 12
- 3) ... check from book ... 0.493

(3) The Bayes Formula

$$P(A_i), P(B|A_i), \dots$$

$$P(A_i|B) = \frac{P(A_i|B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)}$$

↳ The backward conditional prob.

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{\infty} P(A_j)P(B|A_j)}$$

Back to example *

$A_i = \text{"There are } i \text{ black balls selected from Box 1"}$

$$P(A_0|B) = \frac{P(A_0)P(B|A_0)}{P(B)}$$

We know the outcome of last event, find the prob. of first event.

$$P(A_0|B) < P(A_2|B)$$

~~A~~ A blood test is 95% effective in detecting the disease.
False-positive rate of the test: 1%
0.5% of the population has the disease.
What is the probability that someone has the disease
given the result is positive?

D: "Has the disease"

E: "Test is positive"

$$P(D|E) = ? \quad P(D) = 0.005 \quad P(E|D) = 0.95 \quad P(E|\bar{D}) = 0.01$$

$$\Rightarrow P(D|E) = \frac{P(D)P(E|D)}{P(E)} = \frac{P(D)P(E|D)}{P(D)P(E|D) + P(\bar{D})P(E|\bar{D})} = 0.323^{(?)}$$

Oct 6, 2017
Friday

5. Discrete Random Variables and Probability Models

5.1 Random Variables

Example: Roll a die twice $X = \text{"The sum of 2 #'s"}$

$$S = \{(1,1), (1,2), \dots\}$$

$$X = X(\omega), \omega \in S$$

\downarrow
a function of basic events

ω	(1,1)	(1,2)	...	(6,6)
X	2	3	-	12

$$(X=5) : \{\omega | \omega \in S, X(\omega) = 5\} = \{(1,4), (2,3), (3,2), (4,1)\}$$

Possible values of X : 2, 3, ..., 12

$$A_i = \{X=i\}, i = 2, 3, \dots, 12$$

$$\hookrightarrow A_i A_j = \emptyset \quad ; i \neq j$$

$$\hookrightarrow A_2 \cup A_3 \cup \dots \cup A_{12} = S$$

$$\sum_{i=2}^{12} P(X=i) = 1$$

$Y = \text{"The maximum of 2 #'s"}$

w	(1,1)	(1,2)	...	(6,6)
Y	1	2	...	6

Definition: Random variable is a real valued function defined over sample space. $X = X(w), w \in S$

Notes: (i) Random variables (in short R.V or r.v) use capital letters X, Y, Z, X_i, Y_j, Z_3 as variables

(ii) Randomness comes from the fact that the outcome cannot be determined before all trials are completed.

(iii) $\{X \leq x\} = \{w | w \in S, X(w) \leq x\}$ is a meaningful event

(iv) X is a discrete r.v if all possible values of x are either finite or countable infinite

$X: x_1, x_2, x_3, \dots, x_n, \dots$

(v) The probability (mass) function of X : (pmf)

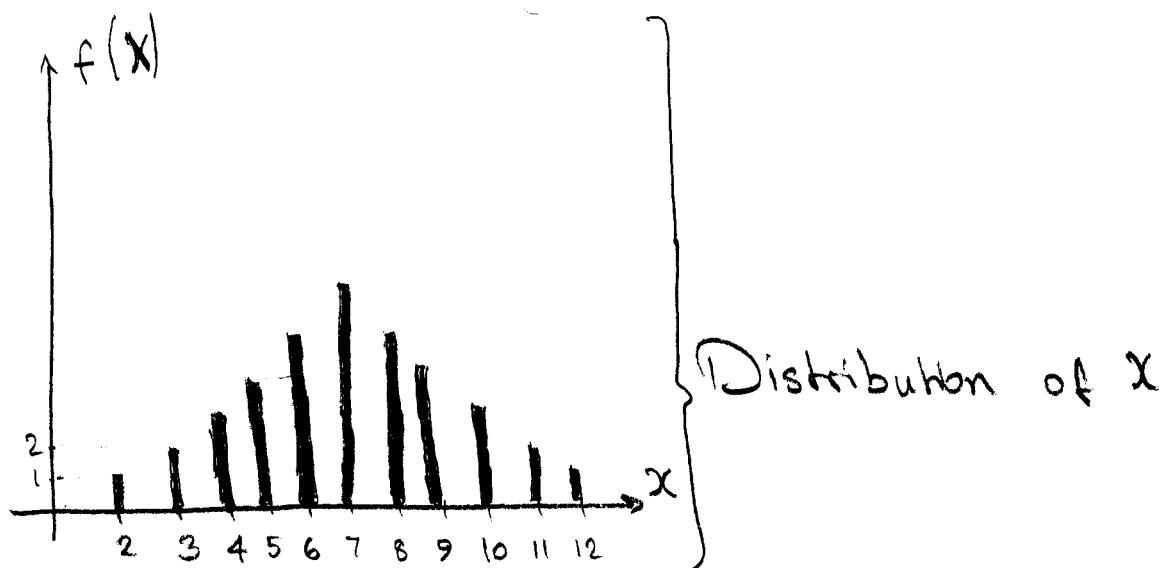
$f(x_i) = P(X = x_i), i = 1, 2, 3, \dots$ generally $f(x) = P(X = x) \forall x$
written as

• Three ways to express $f(x)$:

↳ Table

x	2	3	4	5	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{1}{36}$

↳ The probability bar plot



↳ General math expression

→ Properties of pmf:

$$\hookrightarrow f(x) = P(X=x) \geq 0$$

$$\hookrightarrow \sum_{\text{all } x} f(x) = 1$$

Example: Flip a regular coin n times. $X = \# \text{ of Heads among trials}$

Find the probability function of X .

Q₁: Possible values of X ?

$$\hookrightarrow 0, 1, 2, \dots, n$$

Q₂: What are the probabilities?

$$f(x) = P(X=x) = \frac{\binom{n}{x}}{2^n}$$

$$\text{I} \neq \sum_{x=0}^n f(x) = \sum_{x=0}^n \frac{\binom{n}{x}}{2^n}$$

$$\sum_{x=0}^n \binom{n}{x} = 2^n$$

Recall: The binomial formula

$$(a+b)^n = \sum_{x=0}^n \binom{n}{x} a^x b^{n-x}$$

Example: Three students A, B and C writing exam independently. Their passing probabilities are 0.8, 0.9, 0.85 respectively.

X: "Students who pass the exam"

Q₁: Possible values of X.

$$\hookrightarrow 0, 1, 2, 3$$

Q₂: $f(x) = f(X=x) = ?$

$$\hookrightarrow f(0) = 0.8 \cdot 0.9 \cdot 0.85$$

$$f(1) = 0.8 \cdot 0.1 \cdot 0.15 + 0.2 \cdot 0.9 \cdot 0.15 + 0.2 \cdot 0.1 \cdot 0.85$$

X	0	1	2	3
$f(x)$				

B: "At least two pass the exam"

$$P(B) = P(X \geq 2) = f(2) + f(3)$$

C: "Only Chris passes the exam."

5.2 Bernoulli Trials and Related Distributions

(1) Bernoulli trials, Bernoulli random variable, Bernoulli distribution (Jacob Bernoulli)

- Bernoulli trial: Focus on a particular event

$A = \text{"Success"}$

$\bar{A} = \text{"Failure"}$

$p = P(A)$: The probability of success.

- Bernoulli random variable

$$x = \begin{cases} 1 & \text{if } A \\ 0 & \text{if } \bar{A} \end{cases} \quad \text{The probability function of } X$$

$$f(1) = P(x=1) = P(A) = p$$

$$f(0) = P(x=0) = P(\bar{A}) = 1-p$$

X	0	1
f(x)	1-p	p

$$f(x) = p^x (1-p)^{1-x} \Rightarrow f(1) = p \\ \Rightarrow f(0) = 1-p$$

Bernoulli distribution

- Bernoulli sequence:

↳ Repeat the Bernoulli trial many times

↳ Results from different trials are independent

↳ The success probability, p , remains the same.

(2) Binomial Distribution

→ Bernoulli trial; A; $p = P(A)$

→ Repeat the trial n times.

→ Let X be the number of successes among the n trials.

Find the probability function of X :

$$X: 0, 1, 2, \dots, n$$

$$f(x) = f(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

~~$\sum_{x=0}^n$~~ $n=5$

$$P(X=3)$$

A_i : "ith trial is success"

A_1, A_2, \dots are independent

$$P(A_i) = p$$

$$P(X=3) = \binom{5}{3} p^3 (1-p)^2$$

in 5 trials 3 successes
reserve seat 2 failures

$X \sim \text{Binomial}(n, p)$ (where n is the number of trials
(X follows binomial distribution) and p is probability of success)

$$\sum_{x=0}^n f(x) = \sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x} \stackrel{?}{=} 1$$

Recall: Binomial Expansion

$$(p+q)^n = \sum_{x=0}^n \binom{n}{x} p^x q^{n-x}$$

~~Ex~~

5% of males in a large population is colour blind (CB), independent of each other. Let X be the number of people with CB among 20 randomly selected males.

a) Find prob function of X

b) What's the prob. that at least 2 people are CB.

$$X \sim \text{Binomial}(20, 0.05)$$

$$a) p = 0.05$$

$$1 - p = 0.95 \Rightarrow f(x) = \binom{20}{x} 0.05^x \cdot 0.95^{20-x} \quad x = 0, 1, \dots, 20$$

$$n = 20$$

$$b) P(X \geq 2) = 1 - P(X < 2) = 1 - [f(0) + f(1)]$$

$$= 1 - 0.95^{20} - 20 \cdot 0.05 \cdot 0.95^{19}$$

(3) Geometric Distribution

→ Bernoulli trial; A; p

→ Repeat the trial until 1st success.

→ Let X be the number of failures before 1st success.

The probability function of X :

$$X: 0, 1, 2, \dots \quad (\text{no upper bound})$$

$$f(x) = P(X=x) = (1-p)^x p$$

↙ ↘
 X number of 1st success
 failures

We should have

$$1 \neq \sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} (1-p)^x p = p \cdot \frac{1}{1-(1-p)} = \frac{p}{p} = 1$$

$\overbrace{\quad\quad\quad}$
 $\frac{1}{1-(1-p)}$

Recall: $\sum_{x=0}^{\infty} q^x = 1 + q + q^2 + \dots = \frac{1}{1-q}$

$X \sim \text{Geometric}(p)$

(a) $P(X \geq k) = \sum_{x=k}^{\infty} f(x) = \sum_{x=k}^{\infty} (1-p)^k p = (1-p)^k p \left[1 + (1-p) + (1-p)^2 + \dots \right]$

\nwarrow
at least k failures before success

$$= (1-p)^k$$

(b) $r \leq k: P(r \leq X \leq k) = ?$

(c) X : The discrete waiting time

(d) The memoryless property

$$P(X \geq m+n | X \geq n) = P(X \geq m)$$

Oct 16, 2017

Monday

$$X \sim \text{Geometric}(p)$$

X: Number of failures in a Bernoulli sequence before the 1st success.

$$f(x) = P(X=x) = (1-p)^x p \quad x=0, 1, 2, 3,$$

$$P(X \geq k) = (1-p)^k$$

- Memoryless Property:

$$P(X \geq n+m | X \geq n) = P(X \geq m)$$

$$\text{R.H.S} = (1-p)^m$$

$$\text{L.H.S} = \frac{P(X \geq n+m, X \geq n)}{P(X \geq n)} = \frac{P(X \geq n+m)}{P(X \geq n)} = \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m$$

- What is the probability that X is divisible by 2.

$$\sum_{i=0}^{\infty} P(X=2i) = \sum_{i=0}^{\infty} (1-p)^{2i} \cdot p$$

(4) Negative Binomial Distribution

Bernoulli Sequence

Let X be the number of failures before the k^{th} success.

$X: 0, 1, 2, \dots$

$\begin{matrix} \swarrow & \searrow \\ 0 \text{ failure} & 1 \text{ failure etc.} \end{matrix}$

[e.g.] $k = 3$:

$$f(4) = P(X=4) = \binom{6}{2} p^2 (1-p)^4 p$$

A $\bar{A} \bar{A} A \bar{A} \bar{A}$ A
In 6 trials, 4 failures
 $\underbrace{\quad}_{2 \text{ successes}}$ $\xrightarrow{\quad}$ $\xrightarrow{\quad}$
 3rd success 4th success

$$\underline{f(x) = P(X=x)}$$

$x+k$ trials. Last trial is k^{th} success.

$$\binom{x+k-1}{x} p^{k-1} (1-p)^x p = \underline{\binom{x+k-1}{x} p^k (1-p)^x} = f(x)$$

Show that $\sum_{x=0}^{\infty} f(x) = 1$

$$\sum_{x=0}^{\infty} \binom{x+k-1}{x} p^k (1-p)^x = 1 \Rightarrow p^k \sum_{x=0}^{\infty} \binom{x+k-1}{x} (1-p)^x = 1$$

$$\Rightarrow p^k \cdot (1 - (1-p))^{-k} = 1$$

$$\Rightarrow p^k \cdot p^{-k} = 1$$

$$\Rightarrow 1 = 1$$

- The negative binomial expansion: For any $q \in (0,1)$,

$$(1-q)^{-k} = \sum_{x=0}^{\infty} \binom{x+k-1}{x} q^x$$

~~Ex~~ Suppose 4% of males in a large pop. are colourblind.
Find the prob. of following events.

A: There are 3 CB among a group of 15 randomly selected males.

X = Number of CBs among 15

$X \sim \text{Binomial}(15, 0.04)$

$$P(X=3) = \binom{15}{3} 0.04^3 \cdot 0.96^{12}$$

B: The first CB selected is 15th guy.

X: Number of ~CBs before 1st CB.

X ~ Geometric (0.04)

$$P(X=14) = 0.96^{14} \cdot 0.04$$

C: The 3rd CB is selected by the time 15 non-CBs are already selected.

X: Number of ~CBs before the 3rd CBs.

X ~ Negative Binomial (3, 0.04)

↳ p: probability of success

3 CBs in total

$$P(X=15) = \binom{17}{2} 0.04^2 \cdot 0.96^{15} \cdot 0.04 = \binom{17}{2} 0.04^3 \cdot 0.96^{15}$$

5.3 - Hypergeometric Distribution

Setting: A box contains M red, N-M blue balls, all distinguishable.

Randomly select n balls without replacement.

$$(n \leq M, n \leq N-M)$$

X: Number of red balls in n balls.

$$X: 0, 1, 2, \dots, n$$

$$f(x) = P(X=x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$$

then X ~ Hypergeometric (N, M, n)

Show that $\sum_{x=0}^{M+N} f(x) = 1$ (for hypergeometric distr.)

$$\sum_{x=0}^{M+N} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} = 1 \Rightarrow \sum_{x=0}^{M+N} \binom{M}{x} \binom{N-M}{n-x} = \binom{N}{n}$$

~~Ex~~ 20 students. 7 in STAT, 13 in CS. Randomly select 10 ~~students~~ students.

X: Number of STAT ~~students~~ ^{students}

Find the prob function of X.

$$X = 0, 1, \dots, 7$$

$$f(x) = P(X=x) = \frac{\binom{7}{x} \binom{13}{10-x}}{\binom{20}{10}}$$

At least 2 -

Oct 18, 2017

Wednesday

5.4 Poisson Distribution

- Observational studies; events happen over time and space.

Events follow a poisson process if:

(1) Independence

↳ Events in two non-overlap time intervals are indep.

(2) Individuality

↳ For a small time interval $[t, t + \Delta t]$

$$P(\text{Two or more events in } [t, t + \Delta t]) = o(\Delta t)$$

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$$

(3) Homogeneity

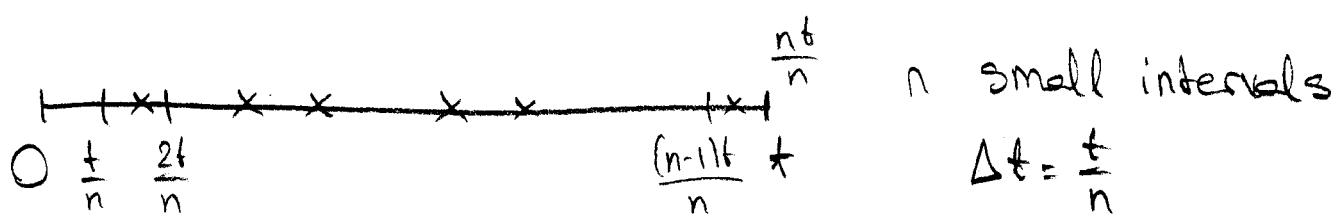
$$P(\text{One event in } [t, t + \Delta t]) = \lambda \cdot \Delta t + o(\Delta t); \lambda: \text{intensity parameter}$$

X : Number of events in $[0, t]$

Find the probability function of X

$X: 0, 1, 2, \dots$

$$f(x) = P(X=x) = ?$$



$$(X=x) = B_1 \cup B_2$$

B_1 : "There are x small intervals each with one event"

B_2 : "At least one small interval with two or more events"

$$\hookrightarrow P(A \cup B) \leq P(A) + P(B)$$

(a) Can show $P(B_2) \rightarrow 0$ as $n \rightarrow +\infty$

$$(b) P(B_1) = \binom{n}{x} [\lambda \Delta t + o(\Delta t)]^x$$

$$= [1 - \lambda \Delta t - o(\Delta t)]^{n-x} = \frac{n(n-1)\dots(n-x+1)!}{x!} \left[\frac{\lambda t}{n} + o\left(\frac{1}{n}\right) \right]^x$$

$$\left[1 - \frac{\lambda t}{n} - o\left(\frac{1}{n}\right) \right]^{n-x}$$

$$= \frac{1}{x!} \left(\frac{n(n-1)\dots(n-x+1)!}{n^x} \right) \left[\lambda t + \frac{o(\frac{1}{n})}{\frac{1}{n}} \right]^x \left[1 - \frac{\lambda t}{n} - o\left(\frac{1}{n}\right) \right]^n \left[1 - \frac{\lambda t}{n} - o\left(\frac{1}{n}\right) \right]^{-x}$$

$$\text{Recall: } \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = e^x$$

As $n \rightarrow \infty$,

$$f(x) = P(X=x) = P(B_1) + P(B_2) = \frac{\mu^x}{x!} e^{-\mu}$$

where $\mu = \lambda t$: The average number of events in $[0, t]$

λ : The average number of events in a unit time

$$\sum_{x=0}^{\infty} f(x) = \left[\sum_{x=0}^{\infty} \frac{\mu^x}{x!} \right] e^{-\mu} = 1$$

$\hookrightarrow e^{\mu}$

Example

Phone calls arrive at a service centre following a Poisson process with an average of 2 calls per min.

a) What's the probability that there are 5 calls in 2 mins.

Let X be the number of events in 2 mins.

$X \sim \text{Poisson}(\mu)$, $\mu = \lambda t$

$$= 2 \cdot 2 = 4$$

- $\begin{matrix} \swarrow \\ \text{cells} \\ \text{per min} \end{matrix}$ $\begin{matrix} \searrow \\ \text{in mins} \end{matrix}$

$$P(X=x) = \frac{\mu^x}{x!} e^{-\mu} \text{ then at } x=5 = \frac{4^5}{5!} e^{-4}$$

(b) What's the probability that there is 1 call in 30 seconds?

Let Y be the number of events in 30 sec.

$$Y \sim \text{Poisson}(\mu), \mu = 2t$$

$$= 2 \cdot \frac{1}{2} = 1$$

$$P(Y=1) = \frac{\mu!}{1!} e^{-\mu} = \underline{\underline{e^{-1}}}$$

(c) A minute is busy if there are at least 5 calls.

What's the prob. a minute is busy?

$Z = *$ of calls in one minute

$$Z \sim \text{Poisson}(\mu), \mu = 2$$

$$P(Z \geq 5) = 1 - P(Z < 4)$$

0.05

(d) A total of 45 minutes will have to be observed to see the 3rd busy minute.

$$\binom{44}{2} p^2 (1-p)^{44-2} \cdot p$$

STAT 240 - Fall 2017

[Announcements](#) > Assignment 3 and Class Schedule

Assignment 3 and Class Schedule

Posted Oct 16, 2017 10:56 PM

Assignment 3 problems are posted in the "Assignments" folder.

As indicated on the Course Outline, there will be **no lectures on Friday October 20 and Monday October 23** (C. Wu is away from campus for the SSC Board Meeting and Statistics Canada Advisory Committee Meeting).

You should use this break to work on the assignment 3 problems. Most of them are doable after the class on Wednesday October 18.

STAT 240 - Fall 2017

[Announcements](#) > Assignment 3 and Class Schedule

Assignment 3 and Class Schedule

Posted Oct 16, 2017 10:56 PM

Assignment 3 problems are posted in the "Assignments" folder.

As indicated on the Course Outline, there will be **no lectures on Friday October 20 and Monday October 23** (C. Wu is away from campus for the SSC Board Meeting and Statistics Canada Advisory Committee Meeting).

You should use this break to work on the assignment 3 problems. Most of them are doable after the class on Wednesday October 18.

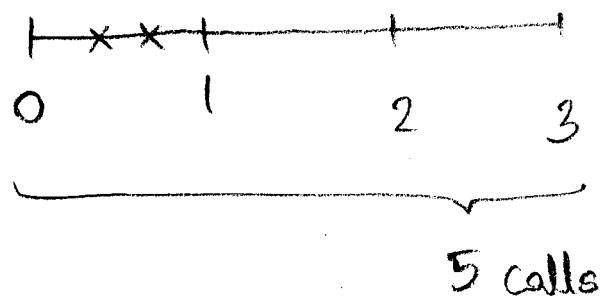
Oct 25, 2011
Wed

~~Ex~~ Phone calls to a service centre.

(c) $P(\text{A minute is busy}) = 0.05$

(d) Minutes are observed one at a time. What's the prob. that a total of 45 mins will have to be observed to see 3rd busy minute?

(e) Suppose there are 5 calls in a 3-min period. What's the prob. that there are 2 calls in first min?



$Y = \# \text{ of calls in 3 mins} \Rightarrow Y = 5 \text{ calls}$

$X_1 = \# \text{ of calls in the 1st min}$

$X_2 = \# \text{ of calls in the last 2 mins}$

X_1, X_2 are ind.

$$Y \sim \text{Poisson}(\mu) \Rightarrow \mu = \lambda t = 2 \cdot 3 = 6$$

$$X_1 \sim \text{Poisson}(\mu_1) \Rightarrow \mu_1 = 2 \cdot 1 = 2 \quad \Rightarrow Y = X_1 + X_2$$

$$X_2 \sim \text{Poisson}(\mu_2) \Rightarrow \mu_2 = 2 \cdot 2 = 4$$

$$P(X_1=2 | Y=5) = \frac{P(X_1=2, Y=5)}{P(Y=5)} = \frac{P(X_1=2, X_2=3)}{P(Y=5)} = \frac{P(X_1=2)P(X_2=3)}{P(Y=5)}$$

$$= \frac{\frac{2^2}{2!} e^{-2} \cdot \frac{4^3}{3!} e^{-4}}{\frac{6^5}{5!} e^{-6}} = \frac{5!}{2!3!} \cdot \frac{2^2 4^3}{6^2 6^3}$$

Recall: $P(x) = \frac{\mu^x}{x!} e^{-\mu}$

$$= \binom{5}{2} \left(\frac{2}{6}\right)^2 \left(\frac{4}{6}\right)^3 \longrightarrow \text{Bernoulli}$$

~~(2)~~ Traffic accidents in K-W area follow Poisson process.
40% of the days have no accident. What's the prob. that there are at least 3 accidents in a day.

$X = \#$ of accidents in a day

$$P(X \geq 3) = 1 - P(X < 3)$$

$$X \sim \text{Poisson } (\mu) \Rightarrow \mu = \lambda t$$

$$= 1 - 0.4 \left(1 + \ln \frac{10}{2} + \frac{(\ln \frac{10}{2})^2}{2} \right)$$

$$P(X=0) = 0.4 \Rightarrow P(X \neq 0) = 0.6$$

$$\text{Recall: } P(X=0) = \frac{\mu^0}{0!} e^{-\mu} = e^{-\mu} = 0.4$$

$$P(X=1) = \frac{\mu^1}{1!} e^{-\mu} = 0.4 \cdot \ln \frac{10}{2}$$

$$P(X=2) = \frac{\mu^2}{2!} e^{-\mu} = 0.4 \cdot \frac{(\ln \frac{10}{2})^2}{2}$$

~~F(x)~~

Visitors arrive at the mall following a poisson process with an average of μ visitors per day. Each visitor has a prob. p to be a window-shopper.

$X = *$ of W-S per day at the mall

$X = 0, 1, 2, \dots$

$$f(x) = P(X=x) = ?$$

$Y = *$ of total visitors per day

$Y \sim \text{Poisson}(\mu)$

$$f(x) = P(X=x) = \sum_{y=0}^{\infty} P(Y=y) P(X=x | Y=y) = \sum_{y=x}^{\infty} P(Y=y) P(X=x | Y=y)$$

$$= \sum_{y=x}^{\infty} \frac{\mu^y}{y!} e^{-\mu} \binom{y}{x} p^x (1-p)^{y-x} = e^{-\mu} p^x (1-p)^x \sum_{y=x}^{\infty} \frac{\mu^y}{y!} \cdot \frac{y! (1-p)^y}{x! (y-x)!} = \frac{e^{-\mu} p^x}{x! (1-p)^x} \sum_{y=x}^{\infty} \frac{\mu^y (1-p)^y}{(y-x)!}$$

$$f(x) = \frac{(\mu p)^x}{x!} e^{-\mu p}; x = 0, 1, 2, \dots$$

$X \sim \text{Poisson}(\mu p)$

$$\text{CDF} = F(x) = P(X \leq x)$$

Oct 27, 2017

Friday

6. Expectation and Variance

↳ 6.1: Expectation of a R.V (mean; average)

$$X \sim f(x)$$

↳ random variable

~~Ex~~

X is the number which is randomly taken from

$$\{3, 5, 5, 7, 7, 7\} \text{ then}$$

x	3	5	7
$f(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

$$E(X) = \text{group average} = (3+5+5+7+7+7)/6$$

$$= 3 \cdot \frac{1}{6} + 5 \cdot \frac{2}{6} + 7 \cdot \frac{3}{6}$$

Defn: X has probability function $f(x)$ if $\sum_x |x| f(x) < \infty$

then the expectation of X is given by

$$E(X) = \sum_x x f(x) = \sum (\text{value})(\text{probability})$$

Notes:

a) $E(X)$ is a fixed numbers.

b) Why $\sum_x |x| f(x) < \infty$



Sequences. Shuffling ^{terms} around give different results.

Harmonic sequence. $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = C$

Absolute convergence

$$E(x) = x_1 f(x_1) + x_2 f(x_2) + \dots \neq x_3 f(x_3) + x_1 f(x_1) + \dots$$

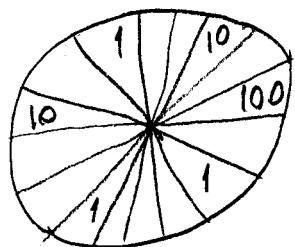
\hookrightarrow sufficient condition

$$X: (-1)^{x-1} \frac{x}{2}, \quad x=1,2,3 \quad f(x) = P\left(X = (-1)^{x-1} \frac{x}{2}\right) = \frac{1}{2^x}$$

$$\sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} (-1)^{x-1} \frac{1}{x}$$

~~Ex~~ Wheel of fortune. The average amount winning.

X = Amount of winning from playing the game once.



X	100	10	1	0
$f(x)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{10}{16}$

$$E(x) = \sum_x x f(x) = 100 \cdot \frac{1}{16} + 10 \cdot \frac{2}{16} + 1 \cdot \frac{3}{16} + 0 \cdot \frac{10}{16}$$

~~Ex~~ Cost effective strategy in medical tests.

Setting: $D = \text{"Disease"}; P(D) = 0.05$

Blood test; cost; \$2 per test

10 samples from 10 diff. people

(i) Method 1: Test the samples one-at-a-time.

$$\hookrightarrow 2 \cdot 10 = \$20$$

(ii) Method 2: The pooled test: Take a half sample from each of the 10 samples. If negative, cost: \$2. If positive, test the 10 samples one-by-one.

X = Number of tests under the pooled strategy.

$$E(X) = 1 \cdot 0.95^{10} + 11 \cdot (1 - 0.95^{10}) \cong 5$$

x	1	11
$f(x)$	0.95^{10}	$1 - 0.95^{10}$

Major steps in applications of expectations:

- (i) Identify the random variable X involved
- (ii) Find the probability function of X
- (iii) Compute $E(X)$.

6.2. Expectation of commonly used R.Vs

* 1) Bernoulli (p)
$$\begin{array}{c|cc} x & 0 & 1 \\ \hline f(x) & 1-p & p \end{array} \Rightarrow E(X) = 0 \cdot (1-p) + 1 \cdot p = p$$

2) Binomial(n, p) $\Rightarrow f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x=0, 1, 2, \dots$

$$\begin{aligned} E(X) &= \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^n x \binom{n}{x} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^n \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)} = \underline{\underline{np}} = E(X) \text{ when } X \sim \text{Binomial}(n, p) \end{aligned}$$

$n=100, p=0.3$ then $E(X)=30$

similar to frequentist interp.

Oct 30, 2017
Monday

$$E(x) = \sum_x x f(x)$$

$$(1) x \sim \text{Bernoulli}(p) : E(x) = p$$

$$(2) x \sim \text{Binomial}(n, p) : E(x) = np$$

$$(3) x \sim \text{Geometric}(p) : E(x) = \sum_{x=0}^{\infty} x(1-p)^x p$$

Recall: $\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}$ where $|a| < 1 \rightarrow$ first derivative

$$\frac{1}{(1-a)^2} = \sum_{i=1}^{\infty} i(a)^{i-1} \rightarrow \text{second derivative} \approx \frac{2}{(1-a)^3} = \sum_{i=2}^{\infty} i(i-2)(a)^{i-2}$$

then $E(x) = \sum_{x=1}^{\infty} x(1-p)^{x-1}(1-p)p = (1-p)p \cdot \frac{1}{(1-(1-p))^2} = \frac{1-p}{p} \cdot E(x)$

~~Ex~~ $p=0.2$ then $E(x) = \frac{0.8}{0.2} = 4$

$$(4) x \sim \text{Negative Binomial}$$

x : Number of failures before k^{th} success.

$$E(x) = \sum_{x=0}^{\infty} x \binom{x+k-1}{x} (1-p)^x p^{k-1} p$$

$$\text{Recall: } \binom{x+k-1}{x} = \frac{(x+k-1)(x+k-2) \cdots k}{x!}$$

$$x \binom{x+k-1}{x} = \frac{(x+k-1)(x+k-2) \cdots (k+1)}{(x-1)!} \cdot k$$

$$= \binom{x+k-1}{x-1} \cdot k$$

$$\text{then } E(x) = \sum_{k=0}^{\infty} x \binom{x+k-1}{x-1} (1-p)^k p^{k-1} p = \sum_{k=0}^{\infty} \binom{x+k-1}{x-1} \cdot k (1-p)^{k-1} (1-p)(p)^{k-1} p^k$$

$$= \frac{k(1-p)}{p} \sum_{k=0}^{\infty} \binom{x+k-1}{x-1} (1-p)^{k-1} p^{k+1} = \underline{\underline{\frac{k(1-p)}{p}}}$$

$$(5) X \sim \text{Hypergeometric}, E(x) = \sum_{x=0}^n x \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$(6) X \sim \text{Poisson } (\mu), E(x) = \mu$$

$$E(x) = \sum_{x=0}^{\infty} x \cdot \frac{\mu^x}{x!} e^{-\mu} = \sum_{x=1}^{\infty} \frac{\mu^{x-1} \mu}{(x-1)!} e^{-\mu} = \mu e^{-\mu} \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!} = \underline{\underline{\mu}} = E(x)$$

↓
Taylor series expansion

6.3: A general formula on expectations

~~Bx~~ X has probability mass function (pmf)

X	-1	0	1	2
f(x)	0.2	0.1	0.3	0.4

Let $Y = X^2$, $E(Y) = ? \implies Y$ can be 0, 1 and 4

$$\text{Recall: } E(Y) = \sum_y y E(Y=y)$$

$$P(Y=0) = P(X=0) = 0.1$$

$$P(Y=1) = P(X=-1) + P(X=1) = 0.2 + 0.3 = 0.5$$

$$P(Y=4) = P(X=2) = 0.4$$

Then $E(Y) = 0 \cdot 0.1 + 1 \cdot (0.2 + 0.3) + 4 \cdot (0.4) = (-1)^2 \cdot (0.2) + 0^2 \cdot 0.1 + 1^2 \cdot 0.3 + 2^2 \cdot 0.4$

Then $E[g(x)] = \sum_x g(x) f(x) \rightarrow$ true in both discrete
and continuous cases
(the law of unconscious statistician)

$$Y = g(x), E(Y) = \sum_y y P(Y=y) \quad Y = y_1, y_2, \dots \quad P(Y=y_j) = P[g(x) = y_j]$$

(1) Expectation is a linear operator: $E(ax+b) = aE(x) + b$

$$\hookrightarrow E(ax^2 + bx + c) = aE(x^2) + bE(x) + c$$

(2) Expectation does not work with non-linear forms.

$$E(x^2) \neq (E(x))^2$$

$$E\left(\frac{1}{x}\right) \neq \frac{1}{E(x)}$$

6.4 Variance

$$X: \begin{array}{c|ccc} x & -1 & 0 & 1 \\ \hline f_x(x) & 0.1 & 0.8 & 0.1 \end{array} \quad E(X) = 0$$
$$Y: \begin{array}{c|ccccc} y & -2 & -1 & 0 & 1 & 2 \\ \hline f_y(y) & 0.4 & 0.5 & 0.1 & 0.5 & 0.4 \end{array} \quad E(Y) = 0$$

6.4 Variance (continued)

$$f(x) = P(X=x)$$

$$\mathbb{E}[g(x)] = \sum_x g(x)f(x)$$

Definition: Let $\mu = \mathbb{E}(x)$ be the expectation value of X . The variance of X is given by $\text{Var}(x) = \mathbb{E}[(x-\mu)^2]$. The expected square deviation from the mean.

Notation:

(i) $\mu = \mathbb{E}(x)$;

(ii) $\sigma^2 = \text{Var}(x)$;

(iii) $\sigma = \sqrt{\text{Var}(x)}$ The standard deviation of x .

for \downarrow
unit consistency

~~Ex~~ Prove $\$1 = \text{£1}$. $\rightarrow \$1 = \100
 $= \$10 \cdot \$10 \rightarrow \text{wrong}$

Computation of Variance

a) $\mu = \mathbb{E}(x)$; $\sigma^2 = \mathbb{E}[(x-\mu)^2] = \sum_x (x-\mu)^2 f(x)$

b) $\sigma^2 = \text{Var}(x) = \mathbb{E}(x^2) - \mu^2$ where $\mathbb{E}(x^2) = \sum_x x^2 f(x)$

c) Useful formula for discrete random variable

$$\sigma^2 = \mathbb{E}(X(X-1)) + \mu - \mu^2$$

$$\text{Recall: } E[x(x-1)] = \sum_x x(x-1) f(x)$$

$$x^2 - x = x(x-1)$$

~~Ex~~ X has pmf:
$$\begin{array}{c|ccc} x & 0 & 1 & 2 \\ \hline f(x) & 0.3 & 0.4 & 0.3 \end{array} \Rightarrow \mu = E(x) = 1$$

$$\sigma^2 = E(x^2) - \mu^2 = \sum_x x^2 f(x) - \mu^2 = 0 + 0.4 + 4 \cdot 0.3 - 1 = 0.6$$

$$\sigma^2 = E((x-\mu)^2) = \sum_x (x-\mu)^2 f(x) = (-1)^2 \cdot 0.3 + 0^2 + 1^2 \cdot 0.3 = 0.6$$

Variance of commonly used R.V.S:

a) $X \sim \text{Bernoulli}(p)$:
$$\begin{array}{c|cc} x & 0 & 1 \\ \hline f(x) & 1-p & p \end{array}$$

$$\mu = E(x) = p; E(x^2) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$$

$$\sigma^2 = p - \mu^2 = p - p^2 = p(1-p)$$

$$\max(\sigma^2) = 0.5 \cdot 0.5 = 0.25$$

b) $X \sim \text{Binomial}(n, p)$: $\mu = np$

$$\begin{aligned} E[x(x-1)] &= \sum_{x=2}^n x(x-1) \binom{n}{x} p^x (1-p)^{n-x} & \text{Recall: } x(x-1) \binom{n}{x} \\ &= n(n-1) \left(\sum_{k=2}^n \binom{n-2}{x-2} p^{x-2} (1-p)^{(n-2)-(x-2)} \right) p^2 &= n(n-1) \binom{n-2}{x-2} \\ &= p^2 n(n-1) \end{aligned}$$

$$\begin{aligned}\sigma^2 &= E[x(x-1)] + \mu - \mu^2 \\&= n(n-1)p^2 + np - n^2 p^2 \\&= -np^2 + np \\&= np(1-p)\end{aligned}$$

c) $X \sim \text{Geometric}(p)$

$$\frac{1}{1-a} = \sum_{x=0}^{\infty} a^x, \quad |a| < 1 \quad \text{and} \quad \frac{1}{(1-a)^2} = \sum_{x=1}^{\infty} x a^{x-1} \quad \text{and} \quad \frac{2}{(1-a)^3} = \sum_{x=2}^{\infty} x(x-1) a^{x-2}$$

$$\text{Recall: } \mu = E(x) = \frac{1-p}{p} \quad \text{then} \quad E[x(x-1)] = \sum_{x=2}^{\infty} x(x-1)(1-p)^x p$$

$$\begin{aligned}\sigma^2 &= E[x(x-1)] + \mu - \mu^2 \\&= \frac{2(1-p)^2}{p^2} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} \\&= \frac{(1-p)^2}{p^2} + \frac{1-p}{p} \quad \frac{1-p}{p^2}\end{aligned}$$

$$\begin{aligned}&= (1-p)^2 p \cdot \frac{2}{(1-(1-p))^3} \\&= \frac{(1-p)^2 2}{p^2}\end{aligned}$$

① d) $X \sim \text{Poisson}(\mu) : \mu = E(x)$

$$E[x(x-1)] = \sum_{x=2}^{\infty} x(x-1) \frac{\mu^x}{x!} e^{-\mu} = \mu^2$$

$$\sigma^2 = E[x(x-1)] + \mu - \mu^2 = \mu$$

Some useful results on $\text{Var}(x)$

(i) $\text{Var}(x) \geq 0$

(ii) $E(x^2) \geq [E(x)]^2$

(iii) $\text{Var}(x) = 0 \iff P(x=c) = 1$

(iv) $\text{Var}(\alpha x + b) = \alpha^2 \text{Var}(x)$

Nov 3, 2017

Friday

$$E(ax+b) = aE(x) + b$$

$$\text{Var}(ax+b) = a^2 \text{Var}(x)$$

$$\text{Var}(x) = E[(x-\mu)^2]$$

$$\text{L.H.S} = E\left[\left(ax+b - E(ax+b)\right)^2\right] = E\left[a^2\left[x - E(x)\right]^2\right] = a^2 E\left[\left[x - E(x)\right]^2\right]$$

~~Ex~~ $X \sim \text{Poisson}$ and $Y = -3x+2$

$$E(Y) = E(-3x+2) = -3E(x)+2 =$$

$$\text{Var}(Y) = \text{Var}(-3x+2) = 9\text{Var}(x)$$

~~Ex~~ Roll a die. Win \$5 if you roll 5 or 6, otherwise you lose \$2.

Do it 100 times. Y = Amount of games that is won.

$$\text{Find } E(Y), \text{Var}(Y). \quad E(Y) = \sum_y y f(y)$$

X = Number of times you won over the 100 games.

$$(i) X \sim \text{Binomial}(100, \frac{2}{6} = \frac{1}{3})$$

$$(ii) Y = 5 \underbrace{x}_{\text{games won}} - 2 \underbrace{(100-x)}_{\text{games lost}} = 7x - 200$$

$$\text{then } E(Y) = 7E(x) - 200 = 7 \cdot 100 \cdot \frac{1}{3} - 200$$

$$\text{Var}(Y) = 49 \text{Var}(x)$$

~~EY~~ Applications in finance.

X: Return from portfolio I (from investing \$1)

x	1.00	1.05	1.10
$f_x(x)$	0.2	0.6	0.2

Average return: $E(X) = 1.05$;

$E(Y) = 1.05$;

Y: Return from portfolio II

y	0.8	1.05	1.30
$f_y(y)$	0.4	0.2	0.4

Volatility: $\text{Var}(X) = 0.001$

$\text{Var}(Y) = 0.05$

X: Low risk

Y: High risk

Volatility is a measurement of risk.

7. Discrete Multivariable Distributions

7.1 Bivariate distributions

~~EY~~ A box contains 2 red balls, 3 white and 4 black. Randomly select 2 balls without replacement.

X = Number of reds selected

$$f(x,y) \sim \begin{cases} x=0,1,2 \\ y=0,1,2 \end{cases}$$

Y = Number of whites selected

		x		
		0	1	2
y	0	6/36	8/36	1/36
	1	12/36	6/36	0
	2	3/36	0	0

$$f(X < Y) = f(0,1) + f(0,2) + f(1,2) = \frac{23}{36}$$

(i) The joint probability function

$$f(x,y) = P(X=x, Y=y) \quad \text{where} \quad \sum_y \sum_x f(x,y) = 1$$

(b) Marginal distributions

$$f_x(x) = P(X=x)$$

$$f_y(y) = P(Y=y)$$

From the previous question, $f_x(0) = P(X=0) = \frac{\binom{7}{2}}{\binom{9}{2}}$

How to find $f_x(x), f_y(y)$ from $f(x,y)$?

$$f_x(0) = P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) + P(X=0, Y=2)$$

$$f_x(x) = P(X=x) = \sum_y P(X=x, Y=y) \quad \text{then} \quad f_x(x) = \sum_y f(x,y)$$

$$f_y(y) = \sum_x f(x,y)$$

Notes:

(i) $f(x,y)$ uniquely determines $f_x(x)$ and $f_y(y)$

(ii) The other direction is not true. Cannot find $f(x,y)$ from $f_x(x), f_y(y)$

(x,y)		x			(x,y)	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$
$f(x,y)$	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
	1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

Nov 6, 2017

Monday

a) Joint probability function of (X, Y) : $f(x, y)$

		x		
		0	1	2
y	0	6/36	8/36	11/36
	1	12/36	6/36	0
	2	3/36	0	0

$$f(x, y) = P(X=x, Y=y)$$

b) Marginal distributions

		x			y		
		0	1	2	0	1	2
$f_x(x)$	0	$\frac{21}{36}$	$\frac{14}{36}$	$\frac{1}{36}$	0	$\frac{15}{36}$	$\frac{18}{36}$
	1	$\frac{14}{36}$	$\frac{6}{36}$	0	1	$\frac{18}{36}$	$\frac{3}{36}$

$$\begin{aligned} f_x(x) &= P(X=x) = \sum_y P(X=x, Y=y) \\ &= \sum_y f(x, y) \end{aligned}$$

c) Conditional distributions

~~Ex~~

The conditional distribution of X given $Y=1$

$$f_{x|y}(x|1) = P(X=x | Y=1), x=0,1,2 = \frac{P(X=x, Y=1)}{P(Y=1)} = \frac{f(x, 1)}{f_y(1)} = \begin{cases} f_{x|y}(0|1) \\ f_{x|y}(1|1) \\ f_{x|y}(2|1) \end{cases}$$

Ask

d) Independent R.V.S

Two events A and B are independent

$$P(A \cap B) = P(A)P(B)$$

Two R.V.S X and Y are indep. if $P(X=x, Y=y) = P(X=x)P(Y=y) \forall (x, y)$

$$\Leftrightarrow f(x, y) = f_x(x) \cdot f_y(y) \text{ for all } (x, y).$$

* Show X and Y are Not ind. Find (x^*, y^*) such that

Show pairs $f(x^*, y^*) \neq f(x^*)f(y^*)$

$$\text{i.e. } f(0,0) \neq f_x(0) \cdot f_y(0)$$

~~Ex~~ $X \sim \text{Geometric}(p)$

$$Y = X^2 \quad P(X=0, Y=1) = 0$$

$$P(X=0) \neq 0$$

$$P(Y=1) \neq 0$$

e) Distribution of a function of R.V.S

~~Ex~~ $(X, Y) \sim f(x, y)$. Let $T = 2XY - 1$

Q₁ → Possible values of T. $\{-1, 1, 3, 7\}$

Q₂ → Probabilities of T.

$$f_T(-1) = P(T=-1) = f(0,0) + f(0,1) + f(0,2) + f(1,0) + f(2,0)$$

~~Ex~~ $X \sim \text{Binomial}(n, p)$

$Y \sim \text{Binomial}(m, p)$

X and Y are independent

a) Find the distribution of $T = X + Y$

Q₁ → $T = 0, 1, 2, \dots, n+m$

$$\left. \begin{array}{c} 0, t \\ 1, t-1 \\ 2, t-2 \\ \vdots \\ t, 0 \end{array} \right\}$$

$$= \sum_{x=0}^t f(x, t-x)$$

$$= \sum_{x=0}^t f(x, t-x) = \sum_{x=0}^t f_x(x) \cdot f_y(t-x) = \sum_{x=0}^t \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{t-x} p^{t-x} (1-p)^{m-(t-x)}$$

$$= p^t (1-p)^{n+m-t} \sum_{x=0}^t \binom{n}{x} \binom{m}{t-x} = \binom{n+m}{t} p^t (1-p)^{n+m-t}$$

$T \sim \text{Binomial}(n+m, p)$

b) Find the conditional dist. of X given $T=t$

$$f_{X|T}(x|t) = P(X=x | T=t) = \frac{P(X=x, T=t)}{P(T=t)} = \frac{P(X=x, Y=t-x)}{P(T=t)}$$

$$= \frac{\binom{n}{x} p^x (1-p)^{n-x} \binom{m}{t-x} p^{t-x} (1-p)^{m-(t-x)}}{\binom{n+m}{t} p^t (1-p)^{n+m-t}} = \frac{\binom{n}{x} \binom{m}{t-x}}{\binom{n+m}{t}}$$

~~Ex~~

$X \sim \text{Poisson}(\mu_1)$

$Y \sim \text{Poisson}(\mu_2)$

X and Y are indep.

a) Find distribution of $T = X+Y$

b) Find the conditional distribution of X given $T=t$

Nov 8, 2017
Wednesday

7.2 Multinomial Distribution

1) Recall: Binomial Distribution.

Trial: Two out comes; A, \bar{A}

$n \downarrow$ times

$$X_1 = \# A \quad P_1 = P(A); \quad P_2 = P(\bar{A}) \Rightarrow P_1 + P_2 = 1$$

$$X_2 = \# \bar{A}$$

$$X_1 + X_2 = n$$

$$P(X_1=x_1, X_2=x_2) = 0, \text{ if } x_1 + x_2 \neq n$$

$$\text{if } X_1 + X_2 = n \text{ then } P(X_1=x_1, X_2=x_2) = \frac{n!}{x_1! x_2!} P_1^{x_1} P_2^{x_2}$$

$$\binom{n}{x} P^x (1-P)^{n-x} = P(x, z)$$

2) Trinomial Distr.

Trial: A, B, C outcomes

$$P_1 = P(A); \quad P_2 = P(B); \quad P_3 = P(C) \quad P_1 + P_2 + P_3 = 1$$

$$\begin{aligned} &n \text{ times and } X_1 = \# A \\ &\quad X_2 = \# B \quad \Rightarrow \quad X_1 + X_2 + X_3 = n \\ &\quad X_3 = \# C \end{aligned}$$

a) Joint prob. function of (x_1, x_2, x_3)

if $x_1 + x_2 + x_3 \neq n$ then $P(x_1=x_1, x_2=x_2, x_3=x_3) = 0$

else $f(x_1, x_2, x_3) = P(x_1=x_1, x_2=x_2, x_3=x_3) =$

Suppose $n=6$ and $P(x_1=3, x_2=2, x_3=1)$

$$P_1^{x_1} P_2^{x_2} P_3^{x_3} \cdot \frac{6!}{3!2!1!} = \binom{6}{3} \binom{3}{2} \binom{1}{1} P_1^3 P_2^2 P_3^1$$

$\downarrow \downarrow \downarrow$ Seat reserving

$$f(x_1, x_2, x_3) = \frac{n!}{x_1! x_2! x_3!} P_1^{x_1} P_2^{x_2} P_3^{x_3} = \binom{n}{x_1} \binom{n-x_1}{x_2} \binom{n-x_1-x_2}{x_3} P_1^{x_1} P_2^{x_2} P_3^{x_3}$$

Notice $x_3 = n - x_1 - x_2$. Then $0 \leq x_1, x_2 \leq n$ and $0 \leq x_1 + x_2 \leq n$

b) Marginal distribution

$$\begin{aligned} f_{X_1}(x_1) &= P(X_1=x_1) = \sum_{x_2=0}^{n-x_1} f(x_1, x_2) \\ &= \sum_{x_2=0}^{n-x_1} \frac{n!}{x_1! x_2! (n-x_1-x_2)!} P_1^{x_1} P_2^{x_2} P_3^{n-x_1-x_2} \\ &= \frac{n!}{x_1! (n-x_1)!} P_1^{x_1} \sum_{x_2=0}^{n-x_1} \frac{(n-x_1)!}{x_2! (n-x_1-x_2)!} P_2^{x_2} P_3^{n-x_1-x_2} \\ &= \frac{n!}{x_1! (n-x_1)!} P_1^{x_1} (P_2 + P_3)^{n-x_1} \end{aligned}$$

$$= \binom{n}{x_1} P_1^{x_1} (1-P_1)^{n-x_1} \quad \text{where } P_2 + P_3 = 1 - P_1$$

$X_1 \sim \text{Binomial}(n, P_1)$

$X_2 \sim \text{Binomial}(n, P_2)$

$X_3 \sim \text{Binomial}(n, P_3)$

c) Conditional distributions

$$f_{X_1|X_3}(x_1|x_3) = P(X_1=x_1 | X_3=x_3) = \frac{P(X_1=x_1, X_3=x_3)}{P(X_3=x_3)}$$

$$= \frac{\frac{n!}{x_1!(n-x_1-x_3)!} x_3! P_1^{x_1} P_2^{n-x_1-x_2} P_3^{x_3}}{\frac{n!}{x_3!(n-x_3)!} P_3^{x_3} (1-P_3)^{n-x_3}} = \frac{(n-x_3)!}{x_1!(n-x_3-x_1)!} \frac{P_1^{x_1} P_2^{n-x_1-x_2}}{(1-P_3)^{n-x_3}} = \binom{n-x_3}{x_1} \left(\frac{P_1}{P_1+P_2}\right)^{x_1} \left(\frac{P_2}{P_1+P_2}\right)^{n-x_1-x_2}$$

$$X_1|X_3=x_3 \sim \text{Binomial}\left(n-x_3, \frac{P_1}{P_1+P_2}\right)$$

$$X_2|X_1=x_1 \sim \text{Binomial}\left(n-x_1, \frac{P_2}{P_2+P_3}\right)$$

d) Find the probability function of $T = X_1 + X_2$
 $T = 0, 1, 2, \dots, n$

$$f_T(t) = P(T=t) = P(X_1 + X_2 = t)$$

$$= \sum_{x_1=0}^t P(X_1=x_1, X_2=t-x_1)$$

$$T = X_1 + X_2 \sim \text{Binomial}(n, p_1 + p_2)$$

$$W = X_2 + X_3 \sim \text{Binomial}(n, p_2 + p_3)$$

Nov 9, 2017

Thursday

$X \sim \text{Binomial}(n, p)$

- X is the number of successes in n trials.

$$f(x) = f(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\text{Binomial expansion: } (p+1-p)^n = \sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x} \quad | \quad E(x) = np$$

$X \sim \text{Negative Binomial}$

- X is the number of failures before the k^{th} success

$$f(x) = f(X=x) = \binom{x+k-1}{x} p^{k-1} (1-p)^x \quad p \rightarrow \text{finishes } (k^{\text{th}} \text{ success})$$

\nwarrow
 x failures in $x+k-1$ trials

$$\sum_{x=0}^n q^x = \frac{1-q^{n+1}}{1-q}$$

$X \sim \text{Geometric}(p)$

- X is the number of failures before 1st success

$$f(x) = f(X=x) = (1-p)^x p \quad \nwarrow \text{finishes} \quad \sum_{x=0}^{\infty} (1-p)^x p = p \cdot \frac{1}{(1-(1-p))} = 1 \quad | \quad E(x) = \frac{1-p}{p}$$

$X \sim \text{Hypergeometric}(N, M, n)$

- Contains $M, N-M$ things. Select n balls without replacement.

$$f(x) = f(X=x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$\sum_{x=0}^{\infty} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} = 1 \Rightarrow \sum_{x=0}^{\infty} \binom{M}{x} \binom{N-M}{n-x} = \binom{N}{n}$$

$X \sim \text{Poisson}(\mu)$

$$\mu = \lambda t, \quad f(x) = P(X=x) = \frac{\mu^x}{x!} e^{-\mu}$$

\nwarrow
intensity parameter

$$\sum_{x=0}^{\infty} \frac{\mu^x}{x!} e^{-\mu} = e^\mu e^{-\mu} = 1$$

$$E(x) = xf(x) = \mu$$

$$\text{Var}(x) = E((x-\mu)^2) = E(x^2) - 2\mu E(x) + \mu^2 \\ = E(x^2) - \mu^2$$

$$= E(x^2 - x + k) - \mu^2 = E(x^2 - x) + \mu - \mu^2 \\ = E(x(x-1)) + \mu - \mu^2$$

Neg. hyper geom.

$$\frac{y^k}{N^k} \binom{y+k-1}{y} (N-M)^y (M)^{k-y}$$

$$\text{Var}(x) = E(x(x-1)) + \mu - \mu^2$$

$$P = \frac{M}{N}$$

$$\mu = NP$$

$$\cancel{\frac{n(n-1)}{N} \frac{M(M-1)}{N}} \quad np(1-p) \left(\frac{N-n}{N-1} \right)$$

Nov 10, 2017

Friday

Example

Three sprinters A, B, C compete in 10 indep. 100 meter races. The prob. of winning a single race are 0.5 for A, 0.4 for B and 0.1 for C.

X_1 : # races won by A.

X_2 : " B.

X_3 : " C.

$X_1, X_2, X_3 \sim \text{Trinomial}(10, 0.5, 0.4, 0.1)$

a) Prob. the A wins 6, B wins 3 and C wins 1.

$$P(X_1=6, X_2=3, X_3=1) = \frac{10!}{6!3!1!} \cdot (0.5)^6 (0.4)^3 \cdot 0.1$$

b) Prob. the B wins 3 races.

$$P(X_2=3) = \binom{10}{3} 0.4^3 (1-0.4)^7 \rightarrow \text{Binomial}$$

c) If A wins 6, what's the prob. that B wins 3.

$$P(X_2=3 | X_1=6) = \binom{10-6}{3} \left(\frac{0.4}{0.5}\right)^3 \left(\frac{0.1}{0.5}\right)^{4-3}$$

relative frequency, A is out of the picture

d) Prob. that A and B win ^{total of} combined 9 races.

$$P(X_1+X_2=9) = \binom{10}{9} (0.9)^9 (0.1)$$

(3) Multinomial Distribution

Trial has $k \geq 2$ possible outcomes: A_1, A_2, \dots, A_k

$$p_i = P(A_i), i=1,2,\dots,k \quad \text{and} \quad \sum_{i=1}^k p_i = 1$$

→ Repeat the trial n times.

$X_i = \# \text{ of trials with outcome } A_i$.

$$X_1 + X_2 + \dots + X_k = n$$

$(X_1, X_2, \dots, X_k) \sim \text{Multinomial distr. } (n, p_1, p_2, \dots, p_k)$

Suppose $X_1 + X_2 + \dots + X_k = n$, then

$$f(x_1, x_2, \dots, x_n) = P(X_1=x_1, X_2=x_2, \dots, X_k=x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

~~Ex/~~ A class has 100 students. Each student receives a letter grade A,B,C,D,F independently with probability 0.1,0.2,0.3,0.3,0.1

$X_i = \# \text{ of students who get } i^{\text{th}} \text{ grade}; i=A,B,C,D,F$

a) Prob. that there are 20A's, 30B's, 35C's, 10D's, 5F's.

$$P(X_1=20, X_2=30, X_3=35, X_4=10, X_5=5) = \frac{100!}{20! 30! 35! 10! 5!} \cdot 0.1^{20} \cdot 0.2^{30} \cdot 0.3^{35+10} \cdot 0.1^5$$

Small numbers, question is asking for specific numbers (strict equality)

b) Prob. that there are at least 50 A's or B's?

$$P(X_1 + X_2 \geq 50) = \sum_{t=50}^{100} P(X_1 + X_2 = t) = \sum_{t=50}^{100} \binom{100}{t} 0.3^t 0.7^{100-t}$$

c) If there are 20 A's, what's the prob. there are 5 F's?

$$\begin{aligned} P(X_5 = 5 | X_1 = 20) &= \binom{100-20}{5} \left(\frac{0.1}{0.9}\right)^5 \left(\frac{0.8}{0.9}\right)^{75} \\ &= \binom{80}{5} \left(\frac{1}{9}\right)^5 \left(\frac{8}{9}\right)^{75} \end{aligned}$$

d) If there are 20 A's, 30 B's what's the prob. there are

5 F's?

$$P(X_5 = 5 | X_1 = 20, X_2 = 30) = \binom{50}{5} \left(\frac{0.1}{0.7}\right)^5 \left(\frac{0.6}{0.7}\right)^{45}$$

e) If there are 5 F's, what's the prob. that there are 20 As and 30 Bs?

$$P(X_1 = 20, X_2 = 30 | X_5 = 5) = \frac{95!}{20! 30! (95-50)!} \cdot \left(\frac{0.1}{0.9}\right)^{20} \left(\frac{0.2}{0.9}\right)^{30} \left(\frac{0.6}{0.9}\right)^{45}$$

remove the conditional

Multinomial will be on
final exam.

7.3 Expectations involving two or more R.V.s

1) The Law of Uncon Statistician

$$X \sim f(x); E[g(x)] = \sum_x g(x) f(x)$$

$$(X_1, X_2) \sim f(x_1, x_2)$$

$$T = g(X_1, X_2) \quad \text{i.e.: } T = X_1 + X_2 \\ T = X_1 X_2$$

$$E(T) = \sum_t t f_T(t) \Rightarrow E[g(X_1, X_2)] = \sum_{x_1} \sum_{x_2} g(x_1, x_2) f(x_1, x_2)$$

$T = g(X_1, X_2)$; possible values t_1, t_2, \dots

$$f_T(t) = P(T=t) \\ = P(g(X_1, X_2) = t) = \sum_{(x_1, x_2): g(x_1, x_2) = t} f(x_1, x_2)$$

Nov 13, 2017
Monday

$$(1) [x, y] \sim f(x, y)$$

$$\text{E}[g(x, y)] = \sum_x \sum_y g(x, y) f(x, y)$$

(2) Basic properties

a) $\text{E}(x+y) = \text{E}(x) + \text{E}(y)$

$$\text{L.H.S.} = \sum_x \sum_y (x+y) f(x, y)$$

$$= \sum_x \sum_y x f(x, y) + \sum_x \sum_y y f(x, y)$$

$$\sum (\alpha x + \beta y + c) = \alpha \text{E}(x) + \beta \text{E}(y) + c$$

b) Usually, $\text{E}(xy) \neq \text{E}(x)\text{E}(y)$

c) If x and y are indep. then $\text{E}(xy) = \text{E}(x)\text{E}(y)$

$$\begin{aligned} \text{E}(xy) &= \sum_x \sum_y xy f(x, y) = \sum_x \sum_y xy f_x(x) \cdot f_y(y) \\ &= \underbrace{\left(\sum_x x f_x(x) \right)}_{\text{E}(x)} \underbrace{\left(\sum_y y f_y(y) \right)}_{\text{E}(y)} \end{aligned}$$

In general, if x and y are indep, $\text{E}[g(x)h(y)] = \text{E}(g(x)) \cdot \text{E}(h(y))$

$$\text{E}(x^2 y^3) = \text{E}(x^2) \cdot \text{E}(y^3) \quad \underline{\text{but}} \quad \text{E}\left(\frac{x}{y}\right) \neq \frac{\text{E}(x)}{\text{E}(y)}$$

Note: $\text{E}(x \cdot \frac{1}{y}) = \text{E}(x) \cdot \text{E}(\frac{1}{y})$

(3) Covariance

$\text{Var}(X+Y) = \text{Var}(T)$ where $T = X+Y$ then,

$$\begin{aligned}\text{Var}(T) &= E[(T - E(T))^2] = E((X+Y - E(X+Y))^2) = E((X - E(X)) + (Y - E(Y)))^2 \\ &= E((X - E(X))^2) + E((Y - E(Y))^2) + 2\text{Cov}(X, Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)\end{aligned}$$

- Definition: The covariance between X and Y is given by

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[(X - \mu_X)(Y - \mu_Y)] \text{ where } \mu_Y = E(Y)$$

Then a) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

b) $\text{Cov}(X, X) = \text{Var}(X)$

c) $\text{Cov}(X, Y) = E(XY) - \mu_X \mu_Y = E(XY) - E(X)E(Y)$

$$\begin{aligned}\therefore E[(X - \mu_X)(Y - \mu_Y)] &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \\ &= E(XY) - \mu_X \mu_Y\end{aligned}$$

d) If X and Y are ind. then $\text{Cov}(X, Y) = 0$

e) $\text{Cov}(X, Y) = 0$ does not imply that X and Y are ind.

i.e.
$$\begin{array}{c|ccc} x & -1 & 0 & 1 \\ \hline f_x(x) & 0.3 & 0.4 & 0.3 \end{array} \quad Y = X^2 \quad \begin{array}{c|cc} y & 0 & 1 \\ \hline f_y(y) & 0.4 & 0.6 \end{array}$$

$\rightarrow X$ and Y are not ind.

$\rightarrow \text{Cov}(X, Y) = E(XY) - E(X)E(Y) ; E(XY) = E(X^3) \text{ and } E(X) = 0 \text{ then}$

$\text{Cov}(X, Y) = E(X^3) = 0. \therefore E(X, Y) = 0 \not\Rightarrow X \text{ and } Y \text{ are ind.}$

f) Recall: $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$

If X and Y are ind, then $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

In general, $\text{Var}(ax+by+c) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab(\text{Cov}(X, Y))$

(4) Correlation coefficient.

The correlation coefficient $\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}}$

a) ρ is unitless

* b) $|\rho| \leq 1$, $|\rho| = 1 \Leftrightarrow Y = aX + b$. If $a > 0$ then $\rho = 1$
 $a < 0$ then $\rho = -1$

Prove $|\rho| \leq 1$: $|\rho| \leq 1 \Leftrightarrow \frac{|\text{Cov}(X, Y)|}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}} \leq 1$

$$\Leftrightarrow |\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}$$

$$\Leftrightarrow |\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]| \leq \sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}$$

$$\Leftrightarrow (\mathbb{E}[(X - \mu_X)(Y - \mu_Y)])^2 \leq \text{Var}(X) \text{Var}(Y)$$

$$\Leftrightarrow (\mathbb{E}[(X - \mu_X)(Y - \mu_Y)])^2 \leq \mathbb{E}[(X - \mu_X)^2] \mathbb{E}[(Y - \mu_Y)^2]$$

Recall: Cauchy-Schwarz inequality

$$\left(\sum_{i=1}^n x_i y_i \right)^2 \leq \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right)$$

For any X, Y : $[\mathbb{E}(XY)]^2 \leq \mathbb{E}(X^2) \mathbb{E}(Y^2)$

→ Proof: For any t , $\mathbb{E}[(X+tY)^2] \geq 0$

$$E[(X+XY)^2] \geq 0 \Rightarrow E(X^2Y^2 + 2XY + X^2) \geq 0$$

$$\Rightarrow \underbrace{t^2a + tb + c}$$

can be solved with quadratic eqn

Nov 15, 2017
Wed

~~Ex~~
 $(X, Y) \sim$

		x		
		0	1	2
y	1	0.4	0.1	0.1
	2	0.1	0.1	0.2

$$\text{Cov}(XY) = 0.18$$

$$\rho = 0.42$$

~~Ex~~

		x		
		0	1	2
y	1	0.1	0.1	0.4
	2	0.2	0.1	0.1

$$\text{Cov}(XY) = -0.18$$

$$\rho = -0.42$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

x	0	1	2
f(x)	0.5	0.2	0.3

y	1	2
f(y)	0.6	0.4

$$E(X) = ?$$

$$E(Y) = ?$$

$$E(XY) = \sum_x \sum_y xy f(x,y)$$

$$= 0 \cdot 1 \cdot 0.4 + 1 \cdot 1 \cdot 0.1 + \dots$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}} ; \text{Var}(X) = E(X^2) - [E(X)]^2$$

Positive correlation: Large values of X tend to be associated with large values of Y; small values of X tend to be associated with small values of Y.

~~Ex~~

X - Body weight

Y - Body height

$$\rho > 0$$

~~Ex~~

X - STAT 240 grade

Y - Number of hours spent not studying prior to final

$$\rho < 0$$

~~Ex~~ $X \sim \text{Poisson}(3)$

$Y \sim \text{Poisson}(4)$

X and Y are ind.

$$\text{Let } U = X+Y$$

$$V = X-Y$$

$$\rho_{U,V} = ?$$

$$\rho_{U,V} = \frac{\text{Cov}(U,V)}{\sqrt{\text{Var}(U)} \sqrt{\text{Var}(V)}}$$

$$; \quad \text{Var}(U) = \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) \\ = 3 + 4 = 7$$

$$= \frac{-1}{7}$$

$$\text{Var}(V) = \text{Var}(X-Y) = 7$$

$$\text{Cov}(U,V) = \text{Cov}(X+Y, X-Y) = \\ = \text{Cov}(X,X) + \text{Cov}(X,-Y) + \text{Cov}(X,Y) + \text{Cov}(Y,-Y)$$

where $\text{cov}(X,Y) = 0 = \text{cov}(X,-Y) \because X$ and Y are ind.

$$\text{Then } \text{cov}(U,V) = \text{cov}(X,X) + \text{cov}(Y,-Y) \\ = \text{Var}(X) - \text{Var}(Y) \\ = 3 - 4 = -1$$

~~Ex~~ $(X_1, X_2, X_3) \sim \text{Trinomial}(n; p_1, p_2, p_3)$. Find P_{X_1, X_2}

$$f(x_1, x_2) = \frac{n!}{x_1! x_2! (n-x_1-x_2)!} p_1^{x_1} p_2^{x_2} p_3^{n-x_1-x_2} ; \quad 0 \leq x_1+x_2 \leq n$$

Method 1:

$$\text{Cov}(X_1, X_2) = E(X_1 X_2) - \overbrace{E(X_1)}^{\text{np}_1} \overbrace{E(X_2)}^{\text{np}_2}$$

$$\text{Var}(X_1) = np_1(n-p_1)$$

$\because X_1, X_2$ follows binomial

$$\text{Var}(X_2) = np_2(n-p_2)$$

$$\text{and } E(X_1 X_2) = \sum_{x_1} \sum_{x_2} x_1 x_2 f(x_1, x_2)$$

Method 2: Consider $X_1 + X_2$

a) $X_1 + X_2 \sim \text{Binomial}(n, p_1 + p_2)$

then $\text{Var}(X_1 + X_2) = n(p_1 + p_2)(1 - p_1 - p_2)$

~~Recall~~ b) $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2)$

then $n(p_1 + p_2)(1 - p_1 - p_2) = np_1(1 - p_1) + np_2(1 - p_2) + 2\text{Cov}(X_1, X_2)$

$$\Rightarrow n \left((p_1 + p_2)(1 - p_1 - p_2) - p_1(1 - p_1) - p_2(1 - p_2) \right) = 2\text{Cov}(X_1, X_2)$$

$\text{Cov}(X_1, X_2) = -np_1 p_2$

- Generalization to multinomial:

$$(X_1, \dots, X_k) \sim \text{Multinomial}(n; p_1, p_2, \dots, p_k) \Rightarrow \text{Cov}(X_i, X_j) = -n p_i p_j$$

7.4. Expectation and Variance of a Linear Combination of R.V.s

(i) Some general results

$$X_1, X_2, \dots, X_n$$

$$E\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i E(X_i) \quad \text{and} \quad E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

Variances:

$$\text{Var}\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \text{Var}(x_i) + 2 \sum_{i < j} \text{Cov}(x_i, x_j)$$

$$= \sum_{i=1}^n \text{Var}(x_i) + \sum_{i \neq j} \text{Cov}(x_i, x_j)$$

If x_1, \dots, x_n are ind. then $\text{Var}\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \text{Var}(x_i)$

$$\text{Also, } \text{Var}\left(\sum_{i=1}^n c_i x_i\right) = \sum_{i=1}^n c_i^2 \text{Var}(x_i) + 2 \sum_{i < j} c_i c_j \text{Cov}(x_i, x_j)$$

If x_1, \dots, x_n are ind. then $\text{Var}\left(\sum_{i=1}^n c_i x_i\right) = \sum_{i=1}^n c_i^2 \text{Var}(x_i)$

Nov 17, 2017

Friday

② Indicator Variable Techniques

~~Ex~~ $X \sim \text{Binomial}(n, p)$ Find $E(x)$ and $\text{Var}(x)$ $X = \# \text{success in } n \text{ trials}$

Let $X_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ trial} \\ 0 & \text{if the } i^{\text{th}} \text{ trial is a success} \end{cases} \quad i = 1, 2, \dots, n$

$$0) X = X_1 + X_2 + \dots + X_n$$

 $X_i \sim \text{Bernoulli}(p)$

$$E(X_i) = p; \text{Var}(X_i) = p(1-p)$$

 X_1, X_2, \dots, X_n are independent

$$E(x) = E(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^n E(X_i) = np$$

$$\text{Var}(x) = \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) = np(1-p)$$

~~Ex~~ $X \sim \text{Hypergeometric}$ $E(x), \text{Var}(x)$

M red, N-M blue, n without replacement

 $X = \# \text{of red}$

Let $X_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ selection is red} \\ 0 & \text{otherwise} \end{cases}$

$$i = 1, 2, 3, \dots, n$$

a) $X = \sum_{i=1}^n X_i$ X_1, X_2, \dots, X_n are not independent.
Hence without replacement.

b)	X_i	0	1
	$f(X_i)$	$\frac{N-M}{N}$	$\frac{M}{N}$

$$E(X) = E(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^n E(X_i) = n \frac{M}{N}$$

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j)$$

$$\text{Var}(X_i) = \frac{M}{N} \left(1 - \frac{M}{N}\right) \rightarrow \frac{M}{N}$$

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i) E(X_j) \rightarrow \frac{M}{N}$$

$$E(X_i X_j) = P(X_i=1, X_j=1)$$

$$= \frac{M(M-1)}{N(N-1)}$$

$$\text{Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) + 2 \binom{n}{2} \left[\frac{M(M-1)}{N(N-1)} - \frac{M^2}{N^2} \right]$$

~~Ex~~ 10 decoration lights are arranged in a row. Each light can be ON with probability 0.7 or OFF with prob. 0.3, indep. of each other. Two adjacent lights are called a SWITCH if they are ON-OFF or OFF-ON.

$$X = \# \text{ of SWITCHES}$$

* Let $X_i = \begin{cases} 1 & \text{if lights } i \text{ and } i+1 \text{ form a sketch} \\ 0 & \text{otherwise} \end{cases}$

$$i = 1, 2, \dots, 9$$

$$(a) X = X_1 + X_2 + \dots + X_9$$

X_1 and X_2 are not ind.

X_i	0	1
$f(x_i)$	$1-p$	$0.3 \cdot 0.7 + 0.7 \cdot 0.3$

$\Rightarrow p$

X_1 & X_2 are not ind.

X_1 & X_3 are ind. $\Rightarrow X_{ij}$ is lin.

$$\mathbb{E}(X_i) = p ; \text{Var}(X_i) = p(1-p)$$

$$\text{Cov}(X_i, X_j) = 0 \text{ if } |i-j| > 1$$

$$\text{Cov}(X_i, X_{i+1}) = \mathbb{E}(X_i X_{i+1}) - \mathbb{E}(X_i) \mathbb{E}(X_{i+1})$$

$$\mathbb{E}(X_i, X_2) = P(X_i = 1, X_2 = 1)$$

$$\mathbb{E}(X_1, X_2) = P(X_1 = 1, X_2 = 1) = 0.7 \cdot 0.3 \cdot 0.7 + (0.3, 0.7) = 0.103$$

~~Bx~~

$$\sum_{i=1}^9 \mathbb{E}(X_i) = 9p . \text{Var}(X) = \text{Var}\left(\sum_{i=1}^9 X_i\right)$$

$$= \sum_{i=1}^9 \text{Var}(X_i) + 2 \sum_{i=1}^8 \text{Cov}(X_i, X_{i+1})$$

(3)

8. Continuous Probability Distributions

8.1: Basic Concepts

① Probability density function (pdf) $[f(x) = \frac{\text{probability}}{\text{density}}]$

~~Ex~~ X is the body height of randomly selected male adult from a very large (infinite) population

a) Possible values of X. X: cm. (80, 250), (0, ∞)

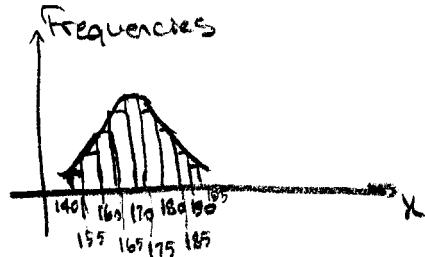
Cover an entire interval.

b) $P(175 \leq X \leq 180) = ?$

Repeated selecting people from the population, to obtain

a selection of people from the desired range.

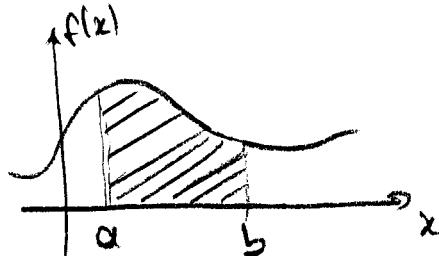
c) Histogram:



d) pdf: X has pdf $f(x)$ if

i) $f(x) \geq 0 \quad \forall x \in (-\infty, \infty)$

ii) For any $a \leq b$, $P(a \leq X \leq b) = \int_a^b f(x) dx$



We say R.V X is continuous with pdf $f(x)$.

Some basic properties

$$\rightarrow P(x=c) = P(c \leq x \leq c) = \int_c^c f(x) dx = 0$$

Note: If $A = \emptyset$ then $P(A) = 0$ but $P(A) = 0$ does not imply $A = \emptyset$

$$\rightarrow P(a < x < b) = P(a \leq x < b) = P(a < x \leq b) = P(a \leq x \leq b) = \int_a^b f(x) dx$$

$$\rightarrow P(X > b) = P(b < x < \infty) = \int_b^\infty f(x) dx$$

$$\rightarrow P(X < a) = P(-\infty < x < a) = \int_{-\infty}^a f(x) dx$$

$$\rightarrow P(-\infty < x < \infty) = \int_{-\infty}^\infty f(x) dx = 1$$

The total area of the density curve is equal to 1.

~~Extp~~

X has pdl

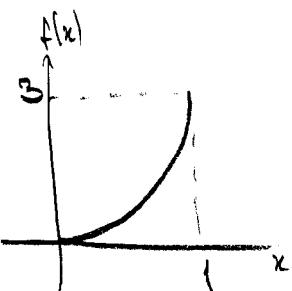
$$f(x) = \begin{cases} kx^2 & \text{if } x \in [0, 1] \\ 0 & \text{if } x \notin [0, 1] \end{cases}$$

redundant

a) $k = ?$

$$f(x) \geq 0; \int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^{\infty} f(x) dx$$

$$\Rightarrow 0 + \frac{k}{3} + 0 = 1 \Rightarrow k = 3$$

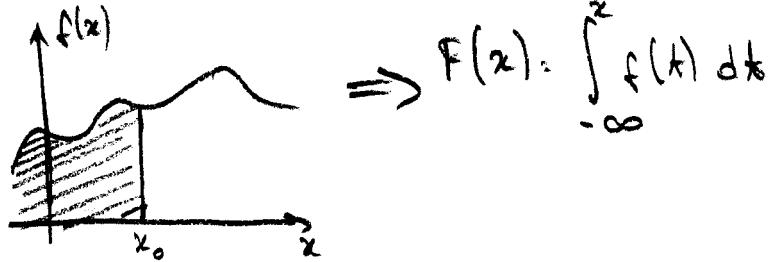


$$b) P\left(\frac{1}{4} < x < \frac{3}{4}\right) = \int_{\frac{1}{4}}^{\frac{3}{4}} 3x^2 dx = \frac{8}{16} = \underline{\underline{\frac{1}{2}}}$$

$$c) P(|x| < \frac{1}{2}) = P\left(-\frac{1}{2} < x < \frac{1}{2}\right) = P(-\frac{1}{2} < x < 0) + P(0 < x < \frac{1}{2}) \\ = 0 + \frac{1}{8} = \underline{\underline{\frac{1}{8}}}$$

② Cumulative Distribution Function (cdf)

$$F(x) = P(x \leq x)$$



Basic properties of cdf:

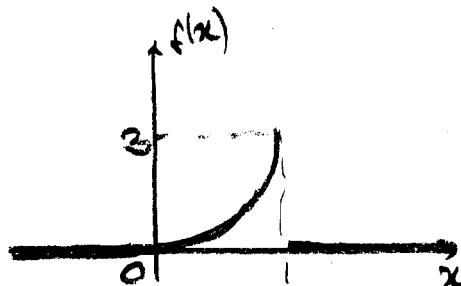
$$(i) 0 \leq F(x) \leq 1$$

$$(ii) F(-\infty) = P(x \leq -\infty) = 0 \text{ and } F(\infty) = P(x \leq \infty) = 1$$

$$(iii) F(x_1) < F(x_2) \iff x_1 < x_2$$

~~x~~ x has pdf $f(x) = 3x^2$. Find the cdf of x .
where $x \in [0, 1]$

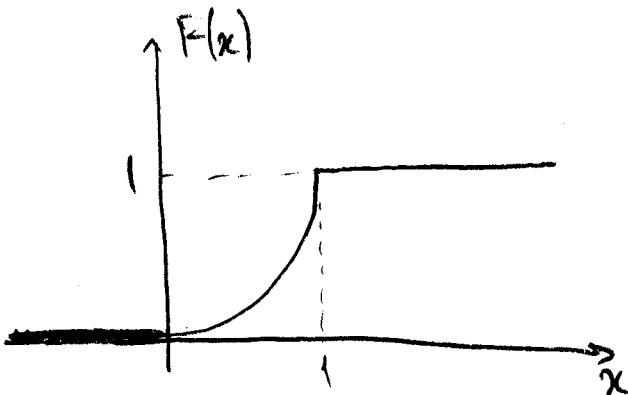
$$a) \text{ if } x \leq 0 \text{ then } F(x) = \int_{-\infty}^x f(t) dt = 0$$



$$b) \text{ if } 0 \leq x \leq 1 \text{ then } F(x) = \int_0^x f(t) dt = x^3$$

$$c) \text{ if } x \geq 1 \text{ then } F(x) = \int_1^x f(t) dt = 1$$

Therefore $F(x) = \begin{cases} 1 & x \geq 1 \\ x^3 & \text{if } x \in (0, 1] \\ 0 & x \leq 0 \end{cases}$



Exercise
for Wednesday

X has pdf

$$f(x) = \begin{cases} kx^2 & x \in [0, 1] \\ k(2-x) & x \in (1, 2] \\ 0 & \text{otherwise} \end{cases}$$

- a) k=?
- b) cdf F(x)

Nov 22, 2017
Wednesday

a) X has pdf:

$$f(x) = \begin{cases} kx^2 & \text{if } x \in [0, 1] \\ k(2-x) & \text{if } x \in (1, 2] \\ 0 & \text{otherwise} \end{cases}$$

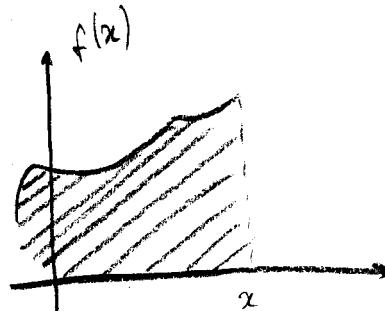
$$k = 6/5$$

b) cdf:

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ (2/5)(x^2) & \text{if } x \in [0, 1] \\ (1/5)(-3x^2 + 12x - 7) & \text{if } x \in (1, 2] \\ 1 & \text{if } x > 2 \\ \text{otherwise} & \end{cases}$$

c) Sketch graph of $f(x), F(x)$

$$P(x) = P(x \leq x) = \int_{-\infty}^x f(t) dt$$



$$P(x \leq b) = F(b) \Rightarrow$$

$$f(a < x \leq b) = F(b) - F(a) \Rightarrow$$

$$f(x > a) = 1 - F(a) \Rightarrow$$

$$\frac{dF(x)}{dx} = f(x) \quad \text{if } f(x) \text{ is continuous at } x$$

(3) Change of variables

X has pdf $f(x)$, $Y = g(X)$ what's the pdf of Y ?

~~Ex~~ X has pdf f

$$f_x(x) = 3x^2, x \in [0, 1]$$

a) $Y = X^2$ Find the pdf of Y .

Range of $Y = [0, 1]$

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

Since range of $X = [0, 1]$ then $F_X(-\sqrt{y}) = 0$

$$\text{Then } F_Y(y) = F_X(\sqrt{y}). \text{ Then } f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(\sqrt{y})}{dy} = f_X(\sqrt{y}) \frac{1}{2\sqrt{y}}$$

b) $Y = -\log X$; $f_Y(y) = ?$

Range of $Y = [0, \infty)$, $y \in [0, \infty)$

$$F_Y(y) = P(Y \leq y) = P(-\log X \leq y)$$

$$= P(X \geq e^{-y}) = 1 - F_X(e^{-y})$$

$$\text{Then } f_Y(y) = \frac{dF_Y(y)}{dy} = f_X(e^{-y}) e^{-y} = \underline{\underline{3e^{-3y}}}$$

$$= \frac{3\sqrt{y}^2}{2\sqrt{y}} \cdot \frac{3\sqrt{y}}{2}$$

where $y \in [0, 1]$

Result: X has pdf $f_x(x)$;

$Y = g(X)$ has a unique inverse.

$$X = g^{-1}(y) = h(y)$$

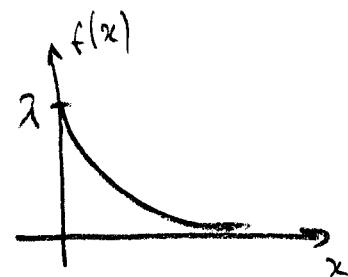
Then $f_y(y) = f_x[h(y)]|h'(y)|$

8.2. Commonly Used Continuous Distributions

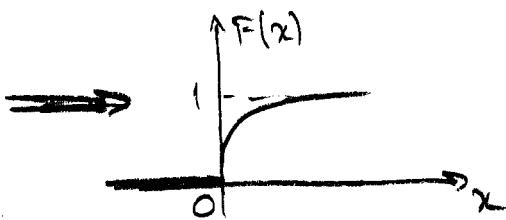
(i) $X \sim \text{Exp}(\lambda)$, The exponential distribution.

a) pdf: $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$ and $\lambda > 0$

$$\hookrightarrow \int_0^\infty \lambda e^{-\lambda x} dx = 1 \quad \Rightarrow$$



b) cdf: $F(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda x} + 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$



$$P(X \leq a) = F(a) = 1 - e^{-\lambda a}$$

$$P(X > a) = 1 - F(a) = e^{-\lambda a}$$

c) X : Life time distribution

d) X : Waiting time.

Poisson process, intensity parameter λ

Let x be the waiting time for the next event.

Range of x : $[0, \infty)$

$$F(x) = P(X \leq x) = 1 - P(X > x) = 1 - P(\text{No events in } [0, x]) \\ = 1 - e^{-\lambda x}$$

e) Memoryless property

$$P(X > t+s | X > s) = \underbrace{P(X > t)}_{e^{-\lambda t}}$$

(1) $X \sim \text{Exp}(\lambda)$

Range of $X: [0, \infty)$

pdf: $f(x) = \lambda e^{-\lambda x}$ $x \geq 0$ and $\lambda > 0$

$$\text{cdf: } F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

(2) Gamma Distributions: A flexible family of distributions over $[0, \infty)$

a) The Gamma function

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx ; \alpha > 0$$

$$\rightarrow \Gamma(\alpha+1) = \alpha \Gamma(\alpha)$$

$$\rightarrow \Gamma(1) = 1 \text{ and } \Gamma(\frac{1}{2}) = \sqrt{\pi}$$

$$\rightarrow \Gamma(n+1) = n!$$

$$\rightarrow \Gamma(\frac{n}{2}) = \frac{n}{2} \cdot \frac{n-2}{2} \cdot \frac{n-4}{2} \cdots \Gamma(\frac{1}{2})$$

b) Gamma Distribution: $X \sim \text{Gamma}(\alpha, \beta)$

→ Range of $X: [0, \infty)$

→ pdf:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad x \geq 0 ; \alpha > 0, \beta > 0$$

and $\int_0^\infty f(x) dx = 1$

3 cases for α : $\alpha < 1, \alpha = 1, \alpha > 1$

When $\alpha = 1 \Rightarrow \text{Gamma}(1, \frac{1}{\beta}) \equiv \text{Exp}(\beta)$

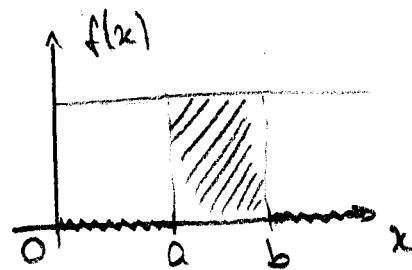
$$\rightarrow \text{cdf}: x > 0, F(x) = \int_0^x f(t) dt = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{x-1} t^{\alpha-1} e^{-\frac{t}{\beta}} dt$$

The incomplete
gamma function

(3) Uniform distributions

a) $U \in [a, b] : a < b$

$$f(x) = \frac{1}{b-a} \quad ; x \in [a, b]$$

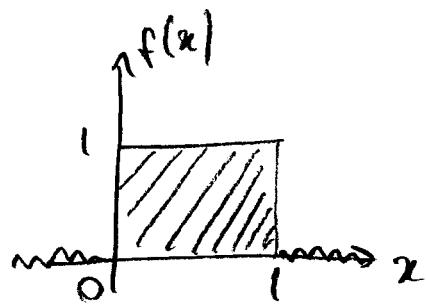


$$\text{Area: } (b-a) \cdot f(x)$$

i.e.: $f(x)=1000 \quad ; x \in [0, 0.001]$

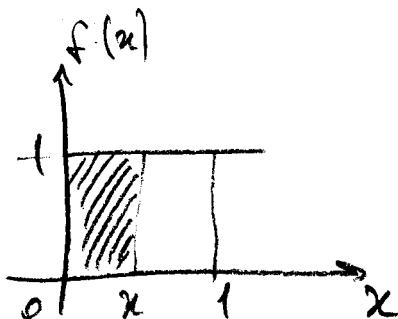
b) $X \sim U[0,1]$

pdf $f: f(x) = 1, x \in [0,1]$



cdf:

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0,1] \\ 1 & \text{if } x > 1 \end{cases}$$



* If $\text{pdf} = 1$ then $X \sim U[0,1]$

~ If cdf follows $F(x)$ as above then $X \sim U[0,1]$

~~Ex~~ X has cdf $F_x(x)$. Let $Y = F_x(X)$. Show $Y \sim U[0,1]$

(Assume $F_x^{-1}(\cdot)$ is unique)

$\rightarrow Y: [0,1]$

$\rightarrow F_Y(y) = y, y \in [0,1]$

- Generating a random numbers X .

Generate $X \sim F(x)$

$$X = F^{-1}(Y) \Rightarrow Y \sim U[0,1]$$

(4) Beta distributions

Range of $X: [0,1]$

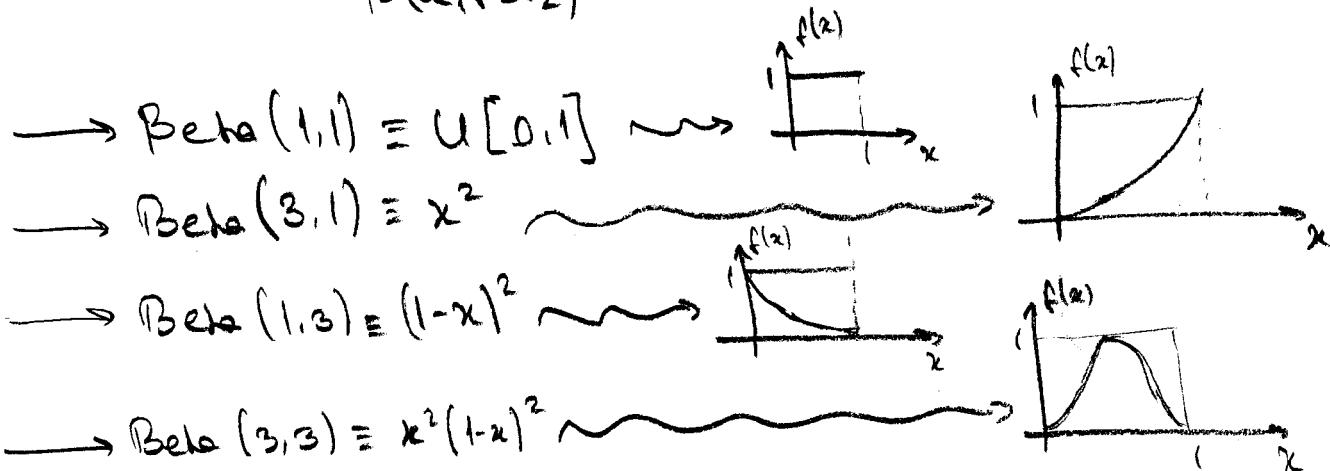
a) The Beta function

$$B(\alpha_1, \alpha_2) = \int_0^1 x^{\alpha_1-1} (1-x)^{\alpha_2-1} dx ; \alpha_1 > 0, \alpha_2 > 0$$

Notice: $B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}$

b) $X \sim \text{Beta}(\alpha_1, \alpha_2)$

$$f(x) = \frac{1}{B(\alpha_1 + \alpha_2)} x^{\alpha_1-1} (1-x)^{\alpha_2-1} ; x \in [0,1]$$



(5) Normal Distributions

~ (0) Standard normal dist.

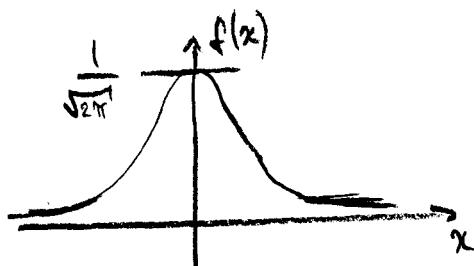
$$\Sigma \sim N(0,1)$$

→ Range of $\Sigma: (-\infty, \infty)$

→ pdf:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in (-\infty, \infty)$$

(i) The Bell shaped density



(ii) $1 = \int_{-\infty}^{\infty} f(x) dx$

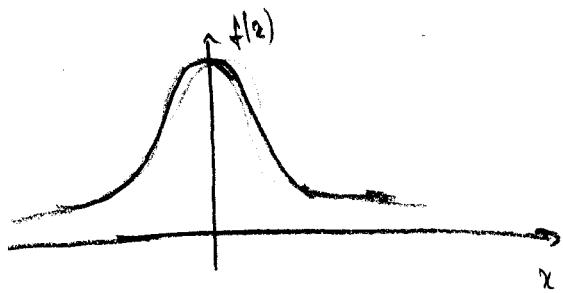
Nov 27, 2017

Monday

(5) Normal Distributions

(a) Standard Normal $Z \sim N(0,1)$

pdf $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in (-\infty, \infty)$



$$a = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$$

$$\sigma^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy$$

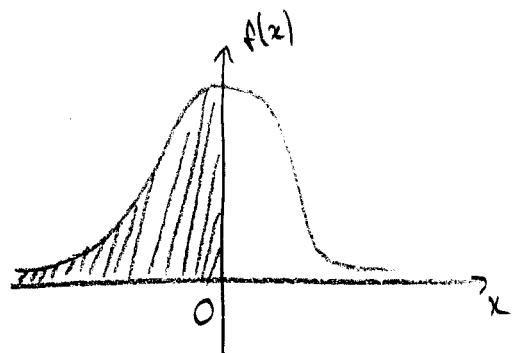
$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

\downarrow
polar coordinates.

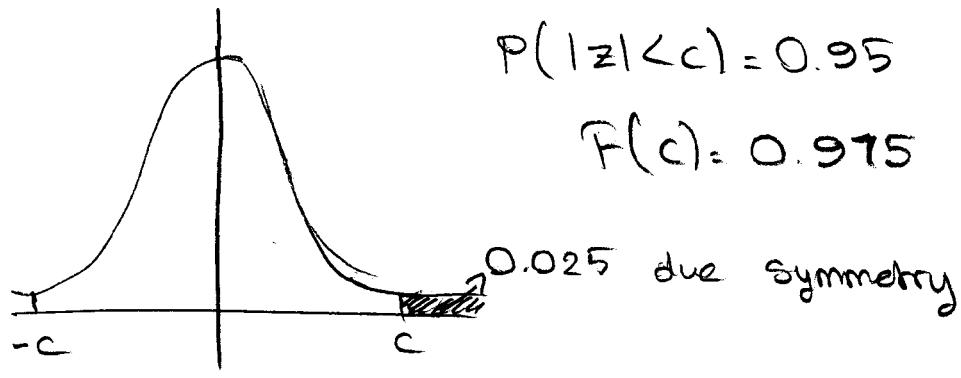
Exercise

Use $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$ to show $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Cdf: $P(Z < z) = F(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$



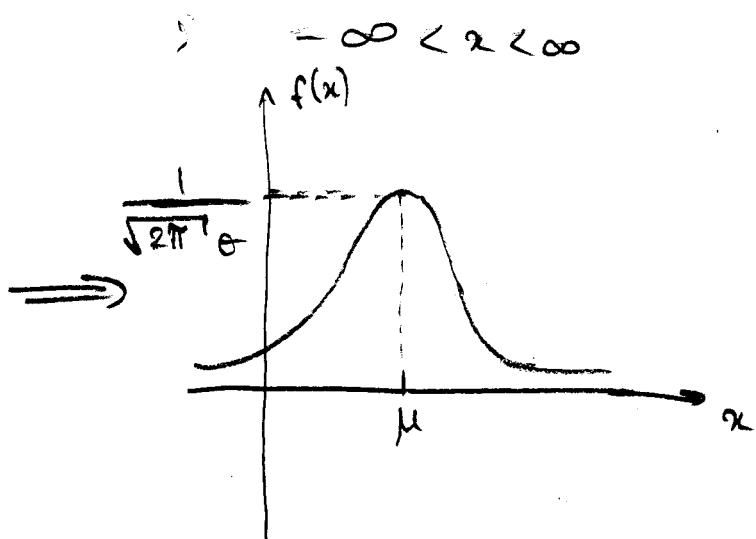
Notice $F(0) = \frac{1}{2}$ because of symmetry



b) $X \sim N(\mu, \sigma^2)$. Range of $X: (-\infty, \infty)$

Pdf: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

$f(x)$ is symmetric at $x=\mu$
 $\max(f(x))$ is when $x=\mu$



Standardization $\rightsquigarrow X \sim N(0,1)$

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{and} \quad Z = \frac{x-\mu}{\sigma} \quad \text{and} \quad z = \frac{x-\mu}{\sigma}$$

then $F_Z(z) = P(Z < z)$

$$= P\left(\frac{x-\mu}{\sigma} \leq z\right)$$

$$= P(X \leq \sigma z + \mu)$$

$$\text{then } F_X(x) = \Phi(x)$$

$$f_z(z) = \frac{d}{dz} F_z(z) = \frac{d}{dz} F_x(\sigma z + \mu)$$

$$= f_x(\sigma z + \mu) \cdot \sigma$$

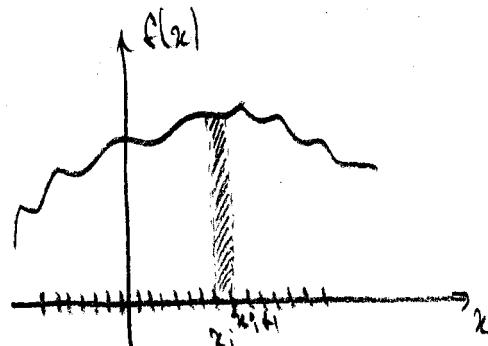
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(z-\mu)^2}$$

8.3 Expectation and Variance

(i) Definition

X is discrete

$$E(X) = \sum_k x_i f(x_i) = \sum_k (\text{value})(\text{prob})$$



The idea of discretization:

- Possible values: $x_1, x_2, \dots, x_i, \dots$

- Probabilities: $f(x_i) \Delta x_i$

$$\sum_i x_i f(x_i) \Delta x_i = \int_{-\infty}^{\infty} f(x) x \, dx$$

X is continuous with pdf $f(x)$, if $\int_{-\infty}^{\infty} |x| f(x) \, dx < \infty$ the expectation of X is given by

$$E(x) = \int_{-\infty}^{\infty} x f(x) \, dx$$

In general,

$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

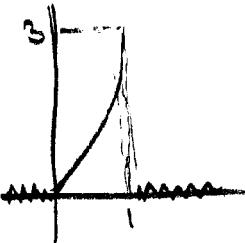
$$\mathbb{E}(x^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

Variance of X : $\text{Var}(x) = \mathbb{E}[(x - \mu)^2]$, $\mu = \mathbb{E}(x)$

$$= \mathbb{E}(x^2) - \mu^2$$

$$\text{Var}(x) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

~~Ex~~
 $f(x) = 3x^2$, $x \in [0, 1]$



$$\mathbb{E}(x) = \int_{-\infty}^{\infty} x f(x) dx = 0 + \int_0^1 x 3x^2 dx + 0 = \frac{3}{4} = \underline{\underline{\mu}}$$

$$\text{Var}(x) = \mathbb{E}(x^2) - \mu^2 = \frac{3}{5} - \frac{9}{16} = \underline{\underline{\frac{3}{80}}}$$

~~Ex~~

$X \sim \text{Exp.}(\lambda)$

$$f(x) = \lambda e^{-\lambda x}, x \geq 0$$

$$\mathbb{E}(x) = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = - \int_0^{\infty} x d e^{-\lambda x} = \frac{1}{\lambda}$$

bigger λ



shorter waiting time

$$\text{Var}(x) = \frac{1}{\lambda^2}$$

Nov 29, 2017

Wednesday

$$\mathbb{E}(x) = \int_{-\infty}^{\infty} x f(x) dx$$

$$\mathbb{E}(x^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$\text{Var}(x) = \mathbb{E}(x^2) - [\mathbb{E}(x)]^2$$

$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

~~Bx~~
 $x \sim \text{Gamma}(\alpha, \beta)$

Find $\mathbb{E}(x), \text{Var}(x)$

$$\begin{aligned}\mathbb{E}(x) &= \int_0^{\infty} x \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\ &= \left(\int_0^{\infty} \frac{1}{\Gamma(\alpha+1)\beta^{\alpha+1}} x^{\alpha+1-1} e^{-\frac{x}{\beta}} dx \right) \cdot \frac{\Gamma(\alpha+1)\beta}{\Gamma(\alpha)}\end{aligned}$$

$$= 1 \cdot \frac{\alpha \Gamma(\alpha) \beta}{\Gamma(\alpha)} = \underline{\underline{\alpha \beta}}$$

~~Ex~~ $X \sim N(0,1)$. $E(x) = ?$ $\text{Var}(x) = ?$

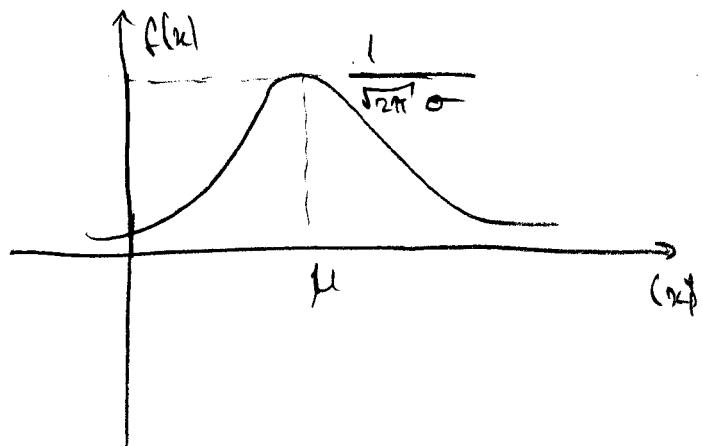
$$E(x) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 \quad \text{and} \quad E(x^2) = 1$$

Symmetric on origin \because odd function

$$f(-x) = -f(x)$$

~~Def~~ $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$



$E(x) = ?$ $\text{Var}(x) = ?$

$$Z = \frac{x-\mu}{\sigma} \sim N(0,1)$$

Standardization

μ : Expectation, average
Central location

σ^2 : Variance

$$X = \sigma Z + \mu \Rightarrow E(x) = \sigma E(Z) + \mu = \mu$$

$$\Rightarrow \text{Var}(x) = \sigma^2 \text{Var}(Z) = \sigma^2$$

* Gamma distribution in assignment.

8.4. Moments and Moment Generating Function (MGF)

(1) Moments: $E(X), E(X^2), \dots$

The r th moment: $\mu_r = E(X^r)$

(2) MGF

$$X \sim f(x) \rightsquigarrow M(t) = E(e^{tx}), t \in (-\infty, \infty)$$

$$\rightarrow \text{pmf } M(t) = \sum_x e^{tx} f(x) \quad \text{includes zero.}$$

$$\rightarrow \text{pdf } M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

~~Ex~~

$$X \sim \begin{array}{c|ccc} x & 0 & 1 & 2 \\ \hline f(x) & 0.6 & 0.3 & 0.1 \end{array} \Rightarrow \underbrace{M(t)}_{\text{pmf}} = e^{0t} \cdot 0.6 + e^{1t} \cdot 0.3 + e^{2t} \cdot 0.1 = 0.6 + e^t \cdot 0.3 + e^{2t} \cdot 0.1 ; t \in (-\infty, \infty)$$

~~Ex~~

$$X \sim \text{Binomial}(n, p)$$

$$M(t) = \sum_{x=0}^n e^{tx} \cdot \underbrace{\binom{n}{x} p^x (1-p)^{n-x}}_{f(x)}$$

$$= \sum_{x=0}^n \binom{n}{x} (e^t p)^x (1-p)^{n-x} = \underline{(e^t p + (1-p))^n}$$

~~Bx~~ $X \sim \text{Gamma}(\alpha, \beta)$

$$M(t) = \int_0^\infty e^{tx} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx$$

$$= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x} e^{-\frac{x(1-\beta t)}{\beta}} dx$$

$$= \frac{1}{(1-\beta t)^\alpha} \int_0^\infty \frac{(1-\beta t)^x}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x} e^{-\frac{x(1-\beta t)}{\beta}} dx$$

$$= \frac{1}{(1-\beta t)^\alpha} \quad \text{where } t < \frac{1}{\beta} \quad \because 1-\beta t > 0 \quad -\beta t > -1$$

~~E_x~~ $X \sim N(0,1)$

$$M(t) = \int_{-\infty}^\infty e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2 + \frac{1}{2}t^2} dx$$

$$= e^{\frac{t^2}{2}} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dx$$

$\underbrace{\qquad\qquad\qquad}_{N(t,1)}$

$$\Rightarrow M(t) = e^{t^2/2}$$

(3) Main properties of MGF

$M(t)$: Laplace transform of $f(x)$

$E(e^{itx})$; i is imaginary number.

Dec 01, 2017
Friday

MGR.

$$M(t) = E(e^{tx}) = \sum_x e^{tx} f(x) \quad \text{or} \quad \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Main properties of MGF:

$$(i) M(0) = E(X); M''(0) = E(X^2)$$

$$M^{(r)}(0) = E(X^r)$$

$$M'(t) = \int_{-\infty}^{\infty} x e^{tx} f(x) dx$$

~~Ex~~ X ~ Geometric(p). Find $E(X^3)$.

$$M(t) = \sum_{x=0}^{\infty} e^{tx} q^x p \quad ; \quad q = 1-p$$

$$= p \sum_{x=0}^{\infty} (qe^t)^x = p \frac{1}{1-e^tq} \quad ; \quad qe^t < 1 \Rightarrow t < -\log q$$

$$(ii) Y = aX + b$$

$$\begin{aligned} M_Y(t) &= e^{bt} M_X(at) \Leftrightarrow E(e^{tY}) = E(e^{t(ax+b)}) \\ &= E(e^{taX}) \cdot e^{tb} \\ &= \underline{\underline{M_X(at) \cdot e^{tb}}} \end{aligned}$$

~~Ex~~ $X \sim N(\mu, \sigma^2)$. Find $M_X(t)$

Let $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Standardization
or normalization

then $M_Z(t) = e^{\frac{t^2}{2}}$ and $X = \sigma Z + \mu$

then $M_X(t) = e^{t\mu} M_Z(\sigma t)$

$$= e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

(iii) X_1 and X_2 are independent

$$Y = X_1 + X_2$$

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

Recall: $M_Y(t) = E(e^{tY})$
 $= E(e^{t(X_1+X_2)})$
 $= E(e^{tX_1} \cdot e^{tX_2})$

then $M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$

(iv) Uniqueness of the Laplace Transform:

X_1, X_2 : cdf $\rightarrow F_1(x), F_2(x)$, MGF $\rightarrow M_1(t), M_2(t)$

$F_1(x) = F_2(x) \iff M_1(t) = M_2(t) \quad ; \quad t \in (-h, h) \rightsquigarrow$ means t can be zero.

~~Ex~~ → Final exam Material

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

$X_2 \sim \text{Binomial}(m, p)$

X_1 and X_2 are independent. Show $\underline{X_1 + X_2 \sim \text{Binomial}(n+m, p)}$

One Approach

$$P(X_1 + X_2 = k) = \sum_{s=0}^k P(X_1 = s, X_2 = k-s)$$

Other Approach

Use MGF.

$$M_{X_1}(t) = (pe^t + 1-p)^n$$

$$M_{X_2}(t) = (pe^t + 1-p)^m$$

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t) = (pe^t + 1-p)^{n+m}$$

[Notice X_1, X_2 share
the same probability]

$M_{X_1+X_2}(t)$ is the MGF of binomial $(n+m, p)$, by the uniqueness
of MGF, we must have $X_1 + X_2 \sim (n+m, p)$

~~Ex~~

$X_1 \sim \text{Gamma}(\alpha_1, \beta)$

$X_2 \sim \text{Gamma}(\alpha_2, \beta)$

X_1 and X_2 are independent. Find the distribution of $X_1 + X_2$.

One Approach

STAT 550?

Change of variable in two dimensions.

Another Approach

$$M_{X_1}(t) = (1 - t\beta)^{-\alpha_1}$$

$$M_{X_2}(t) = (1 - t\beta)^{-\alpha_2}$$

$$M_{X_1+X_2}(t) = (1 - t\beta)^{-(\alpha_1 + \alpha_2)} \Rightarrow X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$$

→ Additivity of distributions.

Also in normal distribution.

~~Ex/~~

Consider a Poisson process with intensity parameter λ .

a) X : The waiting time

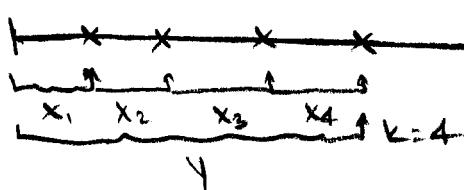
$$X \sim \text{Exp}(\lambda) = \text{Gamma}(1, \frac{1}{\lambda})$$

$$f(x) = \lambda e^{-\lambda x}, x > 0$$

Recall: $\text{Gamma}(\alpha, \beta)$

$$\hookrightarrow x^{\alpha-1} e^{-\frac{x}{\beta}}$$

b) Y : The waiting time for the k^{th} event



X_i : The waiting time between $(i-1)^{\text{th}}$ event and i^{th} event.

$\Rightarrow Y = X_1 + X_2 + \dots + X_k$ and X_i 's are independent of each other because they are not overlapping.

$$\Rightarrow X_i \sim \text{Exp}(\lambda)$$

$$\Rightarrow Y \sim \text{Exp}(k, \frac{1}{\lambda}) \quad \text{and} \quad f(y) = \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-\frac{y}{\beta}}$$

$$= \frac{1}{\Gamma(k) \lambda^k} x^{k-1} e^{-\lambda x}$$

$$= \frac{1}{(k-1)!} \lambda^k x^{k-1} e^{-\lambda x}$$

$$\Gamma(\alpha+1)$$

$$\therefore = \alpha!$$

Dec 04, 2017
Monday

$$\sim X \sim N(0, 1) : M(t) = e^{\frac{t^2}{2}}$$

$$X \sim N(\mu, \sigma^2) : M(t) = e^{\mu t + \frac{1}{2} \sigma^2 t^2}$$

~~Bx~~

$$X_1 \sim N(\mu_1, \sigma_1^2)$$

$$X_2 \sim N(\mu_2, \sigma_2^2)$$

X_1 and X_2 are independent, then

$$C_1 X_1 + C_2 X_2 \sim N\left(C_1 \mu_1 + C_2 \mu_2, C_1^2 \sigma_1^2 + C_2^2 \sigma_2^2\right)$$

Normal distribution

Recall: $Y_1 \sim \text{Binomial}(n, p)$

$Y_2 \sim \text{Binomial}(m, p)$ > independent $Y_1 + Y_2 \sim \text{Binomial}(n+m, p)$

$$M_{X_1}(t) = e^{\mu_1 t + \frac{1}{2} \sigma_1^2 t^2}$$

$$M_{C_1 X_1}(t) = M_{X_1}(C_1 t) = e^{C_1 \mu_1 t + \frac{1}{2} C_1^2 \sigma_1^2 t^2}$$

$$\text{and } M_{X_2}(C_2 t) = e^{C_2 \mu_2 t + \frac{1}{2} C_2^2 \sigma_2^2 t^2}$$

$$\text{Then MGF of } C_1 X_1 + C_2 X_2 = e^{(C_1 \mu_1 + C_2 \mu_2)t + \frac{1}{2} (C_1^2 \sigma_1^2 + C_2^2 \sigma_2^2)t^2}$$

$$\text{Let } Y = C_1 X_1 + C_2 X_2$$

$$\text{then } Y \sim N(E(Y), \text{Var}(Y))$$

$$Y \sim N\left(C_1 \mu_1 + C_2 \mu_2, C_1^2 \sigma_1^2 + C_2^2 \sigma_2^2\right)$$

(1)

~~Bx~~

$$X_i \sim N(\mu_i, \sigma_i^2), i=1,2,\dots,n$$

where X_1, X_2, \dots, X_n are lin. independent

$$c_0 + \sum_{i=1}^n c_i X_i \sim N\left(\left[\sum_{i=0}^n c_i \mu_i\right] + c_0, 0 + \sum_{i=1}^n c_i^2 \sigma_i^2\right)$$

~~Ex~~

X_i are independent and identically distributed (iid)
as $N(\mu, \sigma^2)$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow \text{sample mean}$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$\underbrace{}_{\text{standardization}}$

$$\textcircled{*} \quad \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \equiv \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1)$$

- (*) Is an exact result for normal distributions and also,
- (*) is approximately true for any distribution (discrete; continuous)

Practice Central limit theorem

(2)

(8.5) Central Limit Theorem (CLT)

Y_n : cdf $F_n(y)$; MGF $M_n(t)$
 $(n=1, 2, \dots)$

Y : cdf $F(y)$; MGF $M(t)$

$$\lim_{n \rightarrow \infty} F_n(y) = F(y) \quad \text{for all } y \quad \begin{matrix} \text{(where } F(y) \text{ is)} \\ \text{continuous} \end{matrix}$$

Y_n converges in distribution to $Y \Leftrightarrow \lim_{n \rightarrow \infty} M_n(t) = M(t) ; t \in (-t, t)$

an interval
containing 0.

X_1, X_2, \dots, X_n are iid with an arbitrary distribution $\mu = E(X_i)$

$$\sigma^2 = \text{Var}(X_i)$$

Then, $\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$ converges in distribution to $N(0, 1)$.

Denote W_n for $\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$. Then

$$W_n = \sqrt{n} \frac{1}{n} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} = \frac{1}{\sigma} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} \xrightarrow{} N(0, 1)$$

$$\text{and } M_n(t) \xrightarrow{} M(t) = e^{\frac{t^2}{2}}$$

(3)

$$Y_i = \frac{X_i - \mu}{\sigma} ; \quad i=1,2,\dots,n \quad \text{and } Y_i \text{ are iid.}$$

then $E(Y_i) = 0$ and $\text{Var}(Y_i) = 1$. Recall: $E(Y_i^2) = \text{Var}(Y_i)$

Let $M(t)$ be the MGF of Y_i , then

$$M'(0) = E(Y_i) = 0$$

$$M''(0) = E(Y_i^2) = 1$$

$$M(t) = M(0) + M'(0)t + M''(0)\frac{t^2}{2} + O(t^3)$$

$\underbrace{\hspace{10em}}$ not zero, 0 function

From Taylor Series

$$M(t) = 1 + \frac{t^2}{2} + O(t^3) \quad \text{then MGF of } W_n = \frac{1}{n} \sum_{i=1}^n Y_i$$

$\underbrace{\hspace{10em}}_{[M(t)]^n}$

$$\text{where } [M\left(\frac{t}{n}\right)]^n = \left[1 + \frac{1}{2} \frac{t^2}{n} + O\left(\frac{t^3}{n^{3/2}}\right)\right]^n \xrightarrow{n \rightarrow \infty} e^{\frac{t^2}{2}}$$

END of

STAR 240