University of California, Los Angeles
Department of Statistics

Statistics 13 **Instructor: Nicolas Christou**

Probability models
Some special probability distributions

- **Bernoulli random variable**
  It is a variable that has 2 possible outcomes: "success", or "failure". Success occurs with probability p and failure with probability $1-p$.

- **Geometric probability distribution**

  Suppose that repeated independent Bernoulli trials each one having probability of success p are to be performed. Let X be the number of trials needed until the first success occurs. We say that X follows the geometric probability distribution with parameter p.

  Probability model:

  $$P(X = x) = (1-p)^{x-1}p, \quad x = 1, 2, 3, \cdots$$

Expected value of X:	$E(X) = \frac{1}{p}$
Variance of X:	$\sigma^2 = \frac{1-p}{p^2}$
Standard deviation of X:	$\sigma = \sqrt{\frac{1-p}{p^2}}$

  Example:

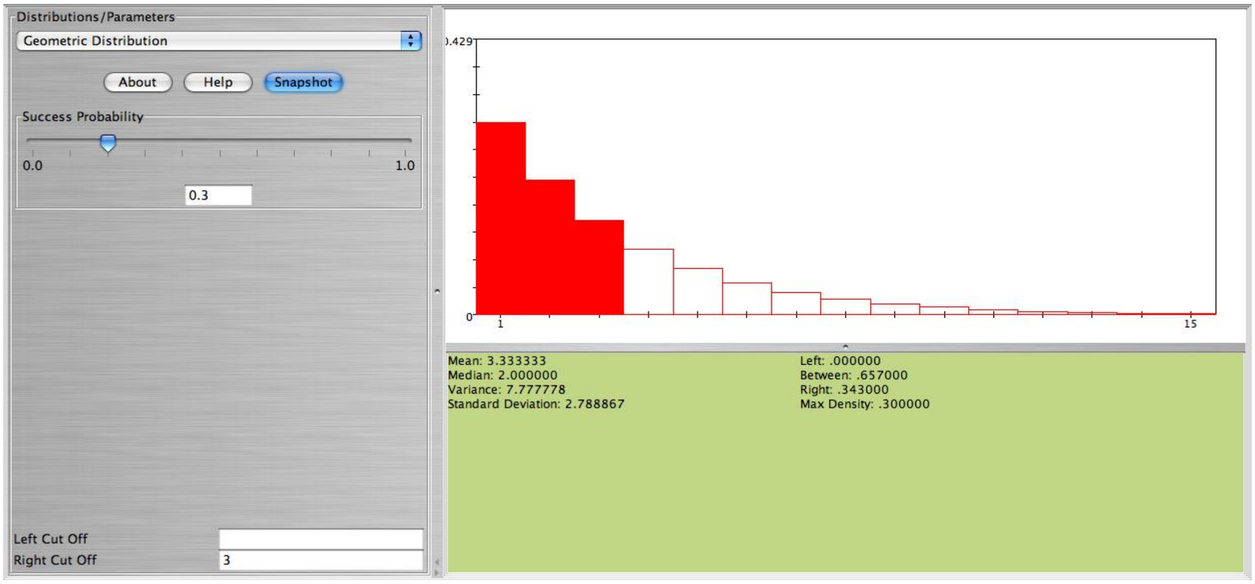
  Patients arrive at a hospital. Assume that 10% of all the patients of this hospital are emergency cases.

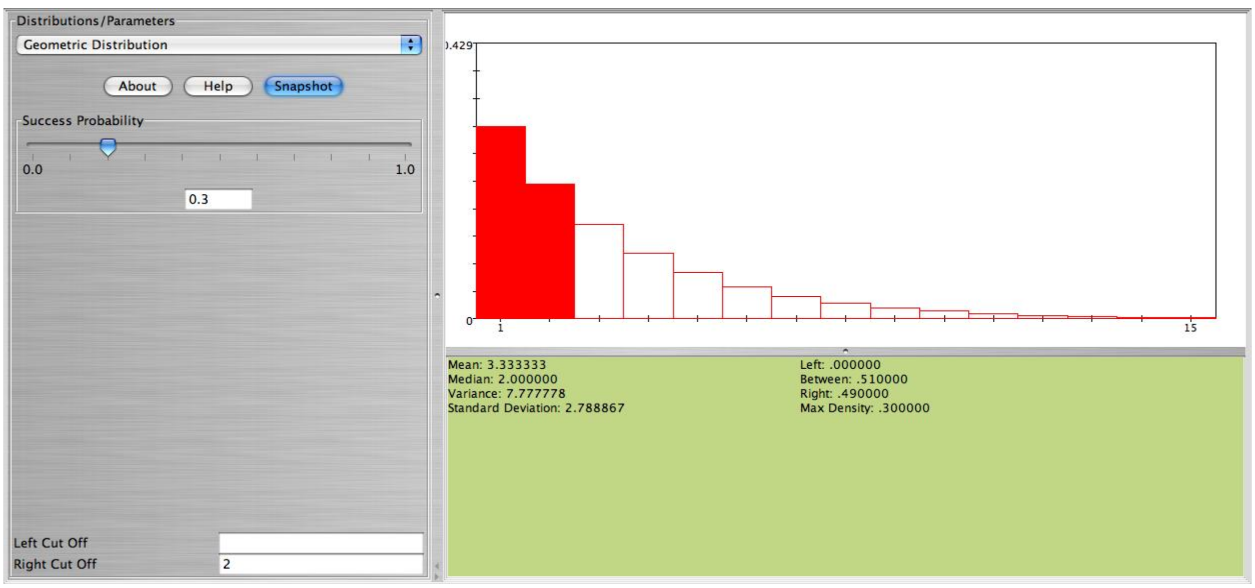
  a. Find the probability that at any given day the 20_{th} patient will be the first emergency case.

  b. On average how many patients must arrive at the hospital to find the first emergency case?

- More on geometric probability distribution $\cdots$
  Repeated Bernoulli trials are performed until the first success occurs. Find the probability that
  - the first success occurs after the k_{th} trial
  - the first success occurs on or after the k_{th} trial
  - the first success occurs before the k_{th} trial
  - the first success occurs on or before the k_{th} trial

Geometric distribution using the Statistics Online Computational Resource (SOCR)



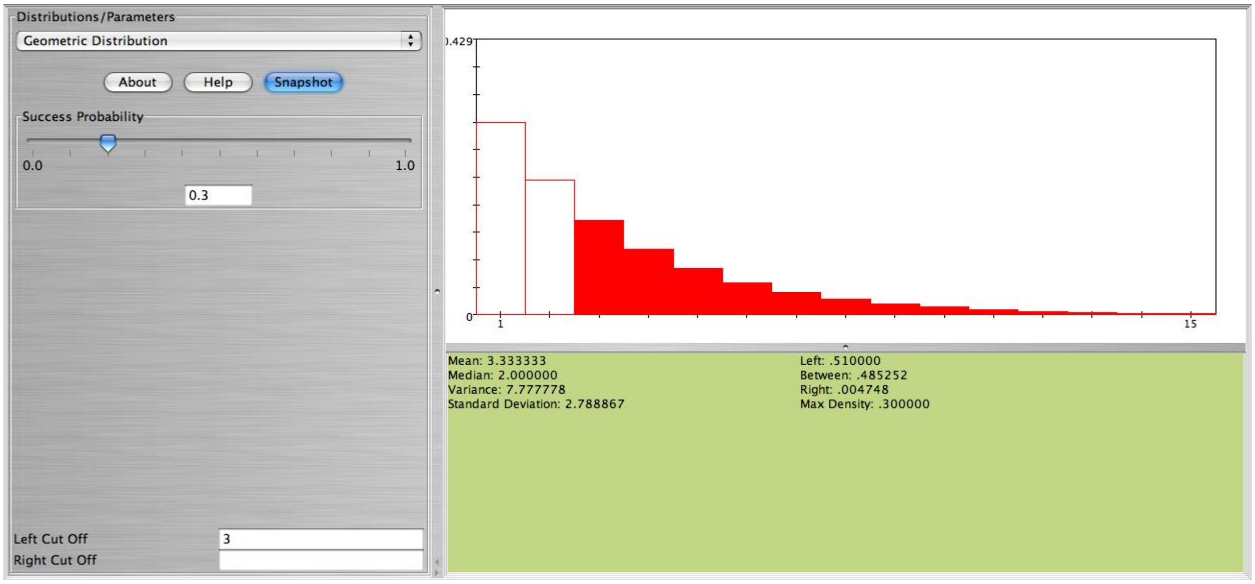


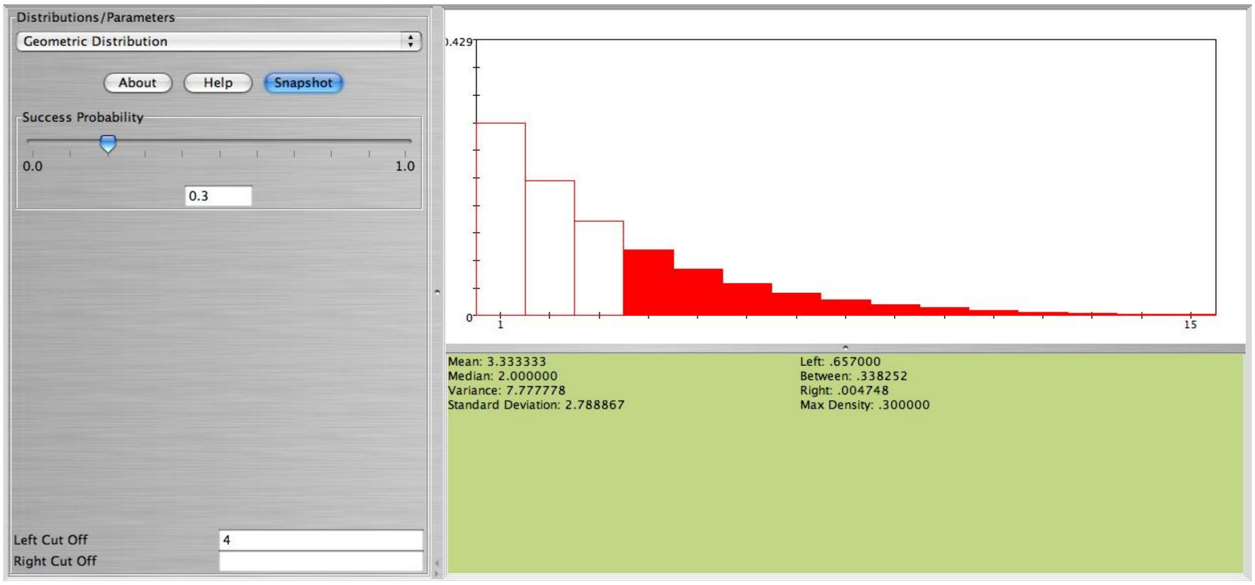
X follows the geometric probability distribution with parameter $p = 0.3$.

a. Write a probability statement in terms of X for the unshaded area of the first applet.

b. Write a probability statement in terms of X for the unshaded area of the second applet.

c. Verify the probabilities in both applets using the formulas $P(X > k) = (1 - p)^k$, etc.



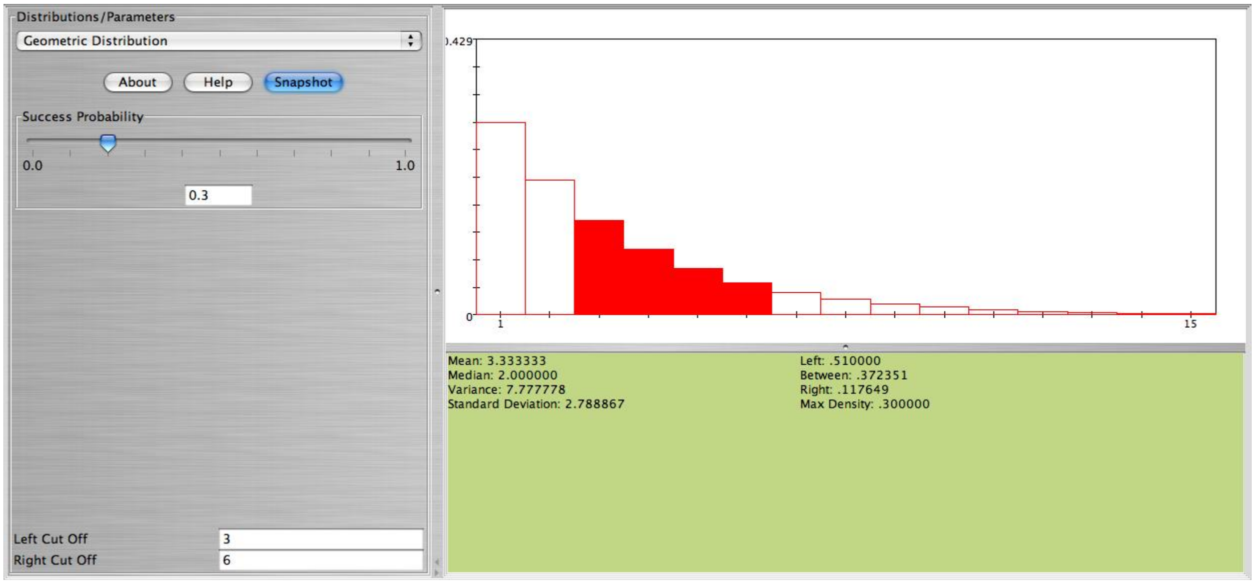


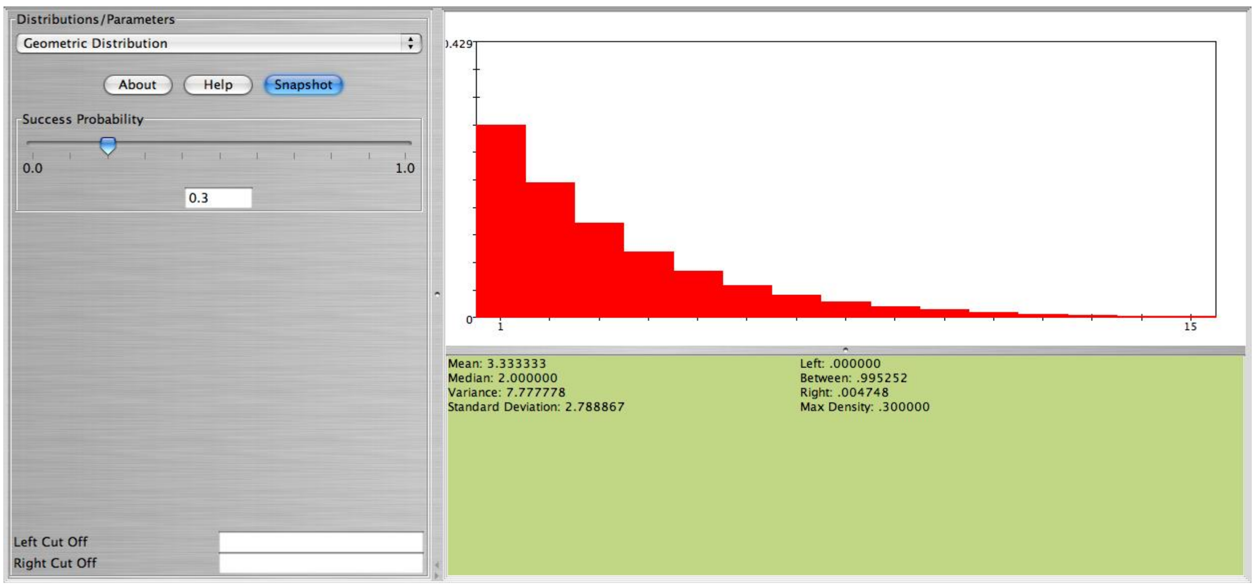
X follows the geometric probability distribution with parameter $p = 0.3$.

a. Write a probability statement in terms of X for the unshaded area of the first applet.

b. Write a probability statement in terms of X for the unshaded area of the second applet.

c. Verify the probabilities in both applets using the formulas $P(X > k) = (1 - p)^k$, etc.





X follows the geometric probability distribution with parameter $p = 0.3$.

a. Write a probability statement in terms of X for the shaded area of the first applet.

b. You want to compute $P(3 \leq X < 9)$. Compute this probability using the formulas, show it on the second applet, and write the Left Cut Off and Right Cut Off points.

- **Binomial probability distribution**
  Suppose that n independent Bernoulli trials each one having probability of success p are to be performed. Let X be the number of successes among the n trials. We say that X follows the binomial probability distribution with parameters n, p.
  Probability model:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, 3, \cdots, n$$

or

$$P(X = x) = nCx \ p^x (1-p)^{n-x}, \quad x = 0, 1, 2, 3, \cdots, n$$

where

$$nCx = \binom{n}{x} = \frac{n!}{(n-x)!x!}$$

Expected value of X:	$E(X) = np$
Variance of X:	$\sigma^2 = np(1-p)$
Standard deviation of X:	$\sigma = \sqrt{np(1-p}$

Example:
Suppose 20 patients arrive at a hospital on any given day. Assume that 10% of all the patients of this hospital are emergency cases.

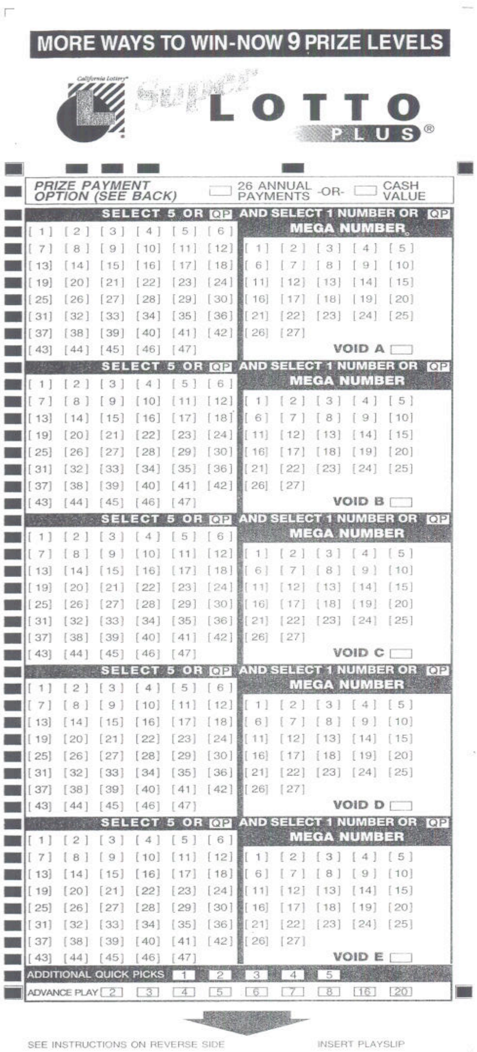
a. Find the probability that exactly 5 of the 20 patients are emergency cases.
b. Find the probability that none of the 20 patients are emergency cases.
c. Find the probability that all 20 patients are emergency cases.
d. Find the probability that at least 4 of the 20 patients are emergency cases.
e. Find the probability that more than 4 of the 20 patients are emergency cases.
f. Find the probability that at most 3 of the 20 patients are emergency cases.
g. Find the probability that less than 3 of the 20 patients are emergency cases.
h. On average how many patients (from the 20) are expected to be emergency case?

- **Hypergeometric probability distribution**
  Select without replacement n from N available items (of which r are labeled as "hot items", and $N - r$ are labeled as "cold items"). Let X be the number of hot items among the n.
  Probability mass function of X:

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Expected value of X: $E(X) = r\frac{n}{N}$
Variance of X: $\sigma^2 = \frac{nr(N-r)(N-n)}{N^2(N-1)} = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}$
Standard deviation of X: $\sigma = \sqrt{\frac{nr(N-r)(N-n)}{N^2(N-1)}}$

California Super Lotto Plus:



Play SuperLOTTO Plus®

- Mark 5 numbers from 1 to 47 and 1 MEGA number from 1 to 27 for each play. Do not use red ink.
- Or mark Quick Pick® to randomly select just the 5 numbers, or only the MEGA number, or both the 5 and the MEGA numbers.
- You may select 1 to 5, 10 or 20 additional Quick Pick® plays on one ticket.
- Play the same numbers for consecutive draws by marking Advance Play®. You may select 2-8,16, 20 or 26 Advance plays.
- Choose how you want to be paid if you win the Jackpot prize. Select either "26 Annual Payments" or "Cash Value".
- 26 Annual Payments: Jackpot paid in 26 annual payments that increase each year.
- Cash Value: Single cash payment, estimated at 40% to 50% of the Announced Jackpot. If you do not choose a Prize Payment Option, the Jackpot will be paid in 26 Annual Payments.
- Your ticket will show your numbers, draw date(s), and payment selection.
- If you make a mistake mark the VOID box. DO NOT ERASE.
- Check your ticket selections for accuracy. If inaccurate, cancel the ticket on the same day of purchase and before pool closure, from the retailer that sold the ticket.
- Your ticket not the playslip is your valid receipt. Sign the back of your ticket to indicate ownership.

HOW TO FILL OUT YOUR PLAY SELECTION

CORRECT INCORRECT

If you make a mistake, mark the VOID box. DO NOT ERASE

PRIZE CATEGORIES AND ODDS

MATCHING NUMBERS	PRIZES	(ODDS 1:)
All 5 of 5 and MEGA	Pari-mutuel*	41,416,353
All 5 of 5	Pari-mutuel*	1,592,937
Any 4 of 5 and MEGA	Pari-mutuel*	197,221
Any 4 of 5	Pari-mutuel*	7,585
Any 3 of 5 and MEGA	Pari-mutuel*	4,810
Any 3 of 5	Pari-mutuel*	185
Any 2 of 5 and MEGA	Pari-mutuel*	361
Any 1of 5 and MEGA	Pari-mutuel*	74
None of 5, only MEGA	Pari-mutuel*	49
Overall Odds of Winning		23

* Prize amounts will vary depending on sales levels and number of winners
On any play, only the highest prize won is awarded

For winning Lottery numbers please log on to
www.calottery.com
Play Responsibly. Gambling Problem Help Line: 1-888-277-3115
For any other Lottery information call 1-800-LOTTERY
Telecommunications Device for the deaf (TDD) 1-800-345-4275

TO CLAIM YOUR PRIZE

- All winning tickets must be claimed within 180 days after the draw in which the prize was won.
- There may be different claiming periods for replays, entries, coupons and promotions. Contact the Lottery for details.
- Winning tickets under $600 per play may be claimed at any participating cashing retailer.
- For winning tickets $600 or more, per play, submit a claim form and your original ticket to the Lottery for payment.
- Federal, State law and Lottery rules and regulations apply. Must be at least 18 years of age to purchase, play and claim.

LO 9/04 B1

Discrete probability distributions - examples

Example 1:
Find the probability that 3 out of 8 plants will survive a frost, given that any such plant will survive a frost with probability of 0.30. Also, find the probability that at least one out of 8 will survive a frost. What is the expected value and standard deviation of the number of plants that survive the frost?

Example 2:
If the probabilities of having a male or female offspring are both 0.50, find the probabilities that a family's fifth child is their first son.

Example 3:
A complex electronic system is built with a certain number of backup components in its subsystem. One subsystem has 4 identical components, each with probability of 0.20 of failing in less than 1000 hours. The subsystem will operate if at least 2 of the 4 components are operating. Assume the components operate independently.

a. Find the probability that exactly 2 of the 4 components last longer than 1000 hours.

b. Find the probability that the subsystem operates longer than 1000 hours.

Example 4:
Suppose that 30% of the applicants for a certain industrial job have advanced training in computer programming. Applicants are interviewd sequentially and are selected at random from the pool. Find the probability that the first applicant having advanced training in computer programming is found on the fifth interview.

Example 5:
Refer to the previous exercise: What is the expected number of applicants who need to be interviewed in order to find the first one with advanced training in computer programming?

Example 6:
A missile protection system consists of n radar sets operating independently, each with probability 0.90 of detecting a missile entering a zone.

a. If $n = 5$ and a missile enters the zone what is the probability that exactly 4 radar sets detect the missile? At least one?

b. How large must n be if we require that the probability of detecting a missile that enters the zone be 0.999?

Example 7:
Construct a probability histogram for the binomial probability distribution for each one of the following: $n = 5, p = 0.1$, $n = 5, p = 0.5$, $n = 5, p = 0.9$. What do you observe? Explain.

Example 8:
On a population of consumers, 60% prefer a certain brand of ice cream. If consumers are randomly selected,

a. what is the probability that exactly 3 people have to be interviewed to encounter the first consumer who prefers this brand of ice cream?

b. what is the probability that at least 3 people have to be interviewed to encounter the first consumer who prefers this brand of ice cream?

Example 9:
The *αlpha* marketing research company employs consumer panels to explore preferences for new products. The current inquiry involves a taste test comparison between a standard brand A vanilla ice cream and a new one of reduced sugar. The company recruits panels of five consumers at a time and determines how many of the five prefer the new product. Their objective is to obtain a panel of five in which all five consumers prefer the new product. The actual probability that an individual will prefer the new product is 0.40.

a. What is the probability that all five consumers in a panel will prefer the new product?

b. What is the exact probability that *αlpha* will go through 60 panels (each of five people) without finding a single panel of five which unanimously prefer the new product?

Example 10:
The *Southland* produce uses the following scheme to assign a quality ranking to incoming shipments of peaches.

Select four peaches at random.

If all four peaches are unbruised, then the shipment is classified A.

If three of the four peaches are unbruised, then select an additional four peaches.

If all of these additional four peaches are unbruised, then the shipment is classified A.

If two or more of these additional four peaches are unbruised, then the shipment is classified B.

If zero or one of these additional four peaches are unbruised, then the shipment is classified C.

If two or fewer of the four peaches are unbruised, then the shipment is classified C.

a. If the proportion of unbruised peaches in a shipment is 0.90, find the probability that the shipment will be classified A? Assume that the shipment is very large so that the probability of selecting an unbruised peach practically remains constant.

b. What is the probability that the shipment will be classified C? Same assumption as in part (a).

Example 11:
A communication system consists of n components, each of which will, independently, function with probability 0.40. The total system will be able to operate effectively if at least one-half of its components function.

a. Would you choose a communication system that consists of $n = 5$ components, or a system that consists of $n = 3$ components?

b. Suppose that a complex communication system comsists of $n = 200$ components. What is the mean and standard deviation of the number of components that will function?

Example 12:
A jury of 6 persons is to be selected at random from a group of 25 potential jurors, of whom 12 are black and 13 are white.

a. What is the probability that the majority of the 6 jurors selected will be black (at least 4 black jurors)?

b. What is the probability that among the 6 jurors selected exactly 3 are black?

c. What is the probability that among the 6 jurors selected no more than 3 will be white?

Example 13:
Use the California State Super Lotto Plus to compute the probability of:

a. Winning the 4_{th} prize (match 4 of 5 but not the mega).

a. Winning the 7_{th} prize (match 2 of 5 and the mega).

Practice problems

Problem 1
Let A and B be events having positive probabilities. State whether each of the following statements is (i) true, (ii) false, (iii) could happen.

a. If A and B are mutually exclusive, then they are independent.

b. If A and B are independent then they are not mutually exclusive.

c. $P(A) = P(B) = 0.6$, and A and B are mutually exclusive.

d. $P(A) = P(B) = 0.6$, and A and B are independent.

e. $P(A) = 0.6, P(B) = 0.3, P(A \cap B) = 0.1$ and $P(A' \cap B') = 0.3$.

Problem 2
An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. Their statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.40, whereas this probability decreases to 0.20 for a non-accident-prone person. Assume that 30% of the population is accident prone.

a. What is the probability that that a new policyholder will have an accident within a year of purchasing a policy?

b. What is the probability that that a new policyholder will not have an accident within a year of purchasing a policy?

c. Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is an accident prone?

Problem 3
Answer the following questions:

a. The probability of receiving a one-pair poker hand is 0.42. What is the probability that a player will receive at least one one-pair hand in 10 games?

b. Compute the expected value and standard deviation of the *maximum* number when two dice are rolled. You will have to find the distribution of the maximum (X) first (i.e. complete the table below).

X	$P(X)$
1	
2	
3	
4	
5	
6	

c. A random variable X has $E(X) = 10$ and $SD(X) = 5$. Find $E(2X^2 + X)$.

d. Two players A, B compete for a prize by drawing a card with replacement from a standard 52-card deck until one receives an ace. What is the probability that player A wins? Note: player A begins the game.

Problem 4
Answer the following questions:

a. If $E(X) = 1$ and $Var(X) = 5$ compute $Var(4 + 3X)$.

b. A box contains 4 green and 5 blue marbles. Two marbles are withdrawn randomly without replacement. If they are the same color, then you win \$1.10. If they are the different colors, then you lose \$1.10. Calculate the expected value of the amount you win.

c. An insurance company must pay a policyholder an amount A if some event E occurs within a year. Suppose the event E will occur with probability p. What should the company charge the customer in order that its expected profit will be 10% of A?

d. An electronic system consists of two components which are connected in series. Each component has its backup component and the system fails if a component and its backup component fails. All the components are independent of each other and they fail with probability 0.10. What is the probability that the system functions?