ACTEX Seminar
Exam P

Written & Presented by
Matt Hassett, ASA, PhD

Remember:

1 **This is a review seminar.**
It assumes that you have already studied probability.

2 **This is an actuarial exam seminar.**
We will focus more on problem solving than proofs.

3 **This is an eight hour seminar.**
You may want to study more material.

Other Study Materials:

- ***Probability for Risk Management***
  (Text and Solutions manual)
  Matt Hassett & Donald Stewart
  ACTEX Publications

- ***ACTEX* Study Guide**
  (SOA Exam P/CAS Exam 1)
  Sam Broverman
  ACTEX Publications

Exam Strategy:

1 **Maximize** the number of questions answered correctly.

2 **Do** the easier problems first.

3 **Don't** spend too much time on one question.

Points to Remember:

- A TV screen holds less content than a blackboard; use your handout pages for overview.
- Algebra and calculus skills are assumed and required.
- Expect “multiple skill problems”.
- Calculators?

5

Probability Rules:

- **Negation Rule:**
  $P(E') = P(\sim E) = 1 - P(E)$
- **Disjunction Rule:**
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

6

Probability Rules, cont.:

- **Definition:**

  Two events A and B are called **mutually exclusive** if $A \cap B = \varnothing$

- ***Addition Rule* for Mutually Exclusive Events:**

  If $A \cap B = \varnothing$, $P(A \cup B) = P(A) + P(B)$

7

Exercise:

The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 35% . Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work.

Determine the probability that a visit to a PCP's office results in both lab work and referral to a specialist.

(A) 0.05 (B) 0.12 (C) 0.18 (D) 0.25 (E) 0.35

8

Solution:

Let L be lab work and S be a visit to a specialist.

$$P\left[\sim\left(L \cup S\right)\right] = 0.35 = 1 - P\left[\left(L \cup S\right)\right]$$

$$P\left(L \cup S\right) = 0.65$$

$$P(S) = 0.30 \text{ and } P(L) = 0.40$$

$$P\left(L \cup S\right) = 0.65 = P(L) + P(S) - P\left(L \cap S\right)$$

$$= 0.40 + 0.30 - P\left(L \cap S\right)$$

$$P\left(L \cap S\right) = 0.05$$

Answer A

Venn Diagrams Can Help:

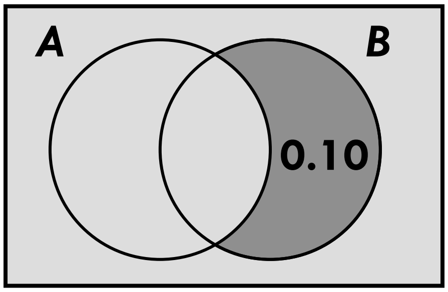
You are given:

$$P\left(A \cup B\right) = 0.7 \text{ and } P\left(A \cup B'\right) = 0.9.$$

Determine $P[A]$.

(A) 0.2 (B) 0.3 (C) 0.4 (D) 0.6 (E) 0.8

Venn Diagrams Can Help:



Unshaded region $P\left(A \cup B'\right) = 0.9.$

Area of the shaded region must be 0.10

The total area of the two circles represents:

$$P\left(A \cup B\right) = 0.7$$

Subtracting the area of the shaded region:

$$P(A) = 0.7 - 0.1 = 0.6$$

Answer D

11

A More Complicated Venn Diagram:

An insurance company has 10,000 policyholders. Each policyholder is classified as young/old; male/female; and married/single.

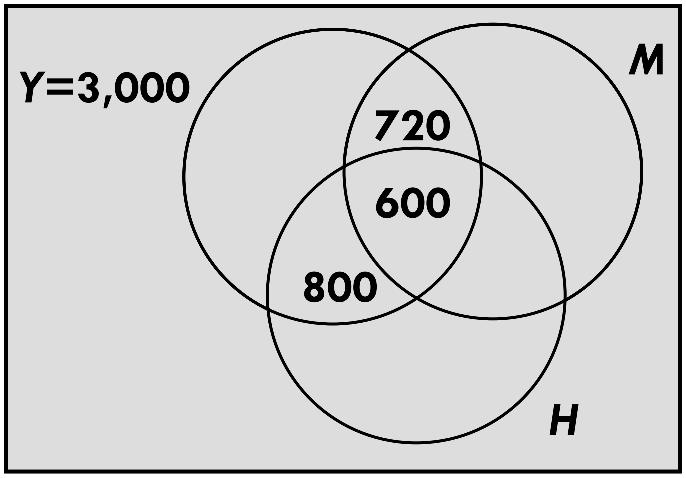
Of these, 3,000 are young, 4,600 are male, and 7,000 are married. They can also be classified as 1,320 young males, 3,010 married males, and 1,400 young married persons. 600 are young married males.

How many policyholders are young, female, and single?

(A) 280 (B) 423 (C) 486 (D) 880 (E) 896

12

A More Complicated Venn Diagram:

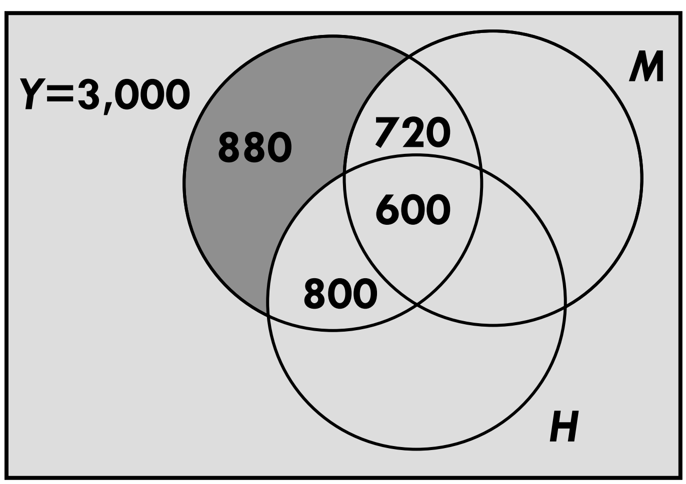


Y = young

M = male

H = married

A More Complicated Venn Diagram:



Y = young

M = male

H = married

$$3,000 - 720 - 600 - 800 = 880$$

Answer D

Some Problems are Trickier:

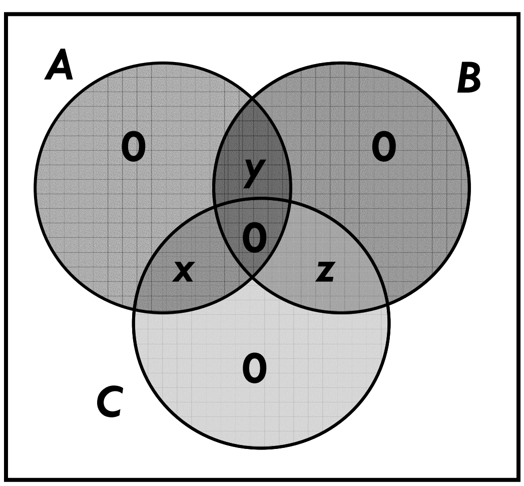
An insurer offers a health plan to the employees of a company. As part of this plan, each employee may choose exactly two of the supplementary coverages *A*, *B*, and C, or may choose no supplementary coverage.

The proportions of the employees that choose coverages *A*, *B*, and C are 1/4, 1/3, and 5/12, respectively.

Determine the probability that a randomly chosen employee will choose no coverage.

(A) 0 (B) 47/144 (C) 1/2 (D) 97/144 (E) 7/9

Trickier Problem, cont.:

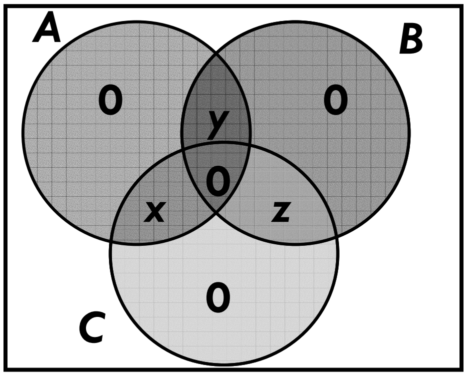


Find

$$P\left[\sim\left(A \cup B \cup C\right)\right]$$
$$= 1 - \left(x + y + z\right)$$

This is a linear system for *x*, *y*, *z*.

Trickier Problem, cont.:



$$P(A) = x + y + 0 = \frac{1}{4}$$

$$P(B) = 0 + y + z = \frac{1}{3}$$

$$P(C) = x + 0 + z = \frac{5}{12}$$

Solution:

$x = 2/12,\ y = 1/12,\ z = 3/12$

$$P\left[\sim\left(A \cup B \cup C\right)\right] = 1 - \left(x + y + z\right) = 1 - \frac{6}{12} = \frac{1}{2}$$

Answer C

17

More Probability Rules:

- **Conditional probability by counting for equally likely outcomes**

$$P(A \mid B) = \frac{n\left(A \cap B\right)}{n\left(B\right)}$$

- **Definition:**

  For any two events A and B, the **conditional probability** of A given B is defined by

$$P(A \mid B) = \frac{P\left(A \cap B\right)}{P\left(B\right)}$$

18

More Probability Rules:

Multiplication Rule for Probability

$$P\left(A \cap B\right) = P(A \mid B)P\left(B\right)$$

Exercise:

A researcher examines the medical records of 937 men and finds that 210 of the men died from causes related to heart disease.

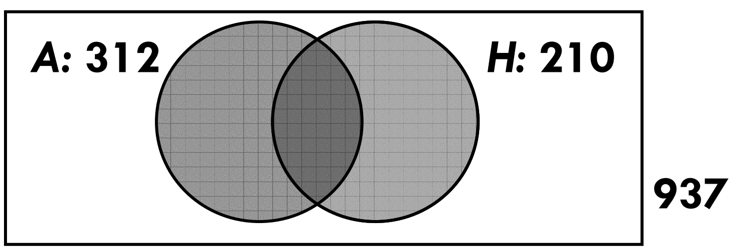
312 of the 937 men had at least one parent who suffered from heart disease, and, of these 312 men, 102 died from causes related to heart disease.

Exercise, cont.:

Find the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.

(A) 0.115
(B) 0.173
(C) 0.224
(D) 0.327
(E) 0.514

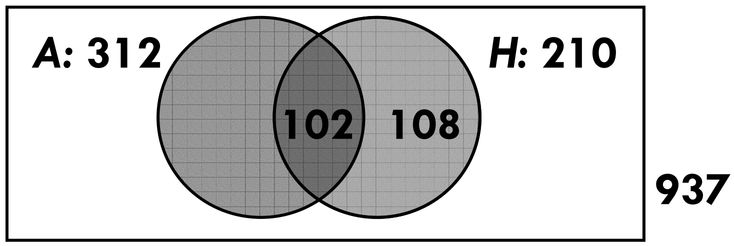
Solution:



A = At least one parent with heart disease
H = Died of causes related to heart disease

Find $P(H \mid \sim A) = \dfrac{n(H \cap \sim A)}{n(\sim A)}$

Solution:

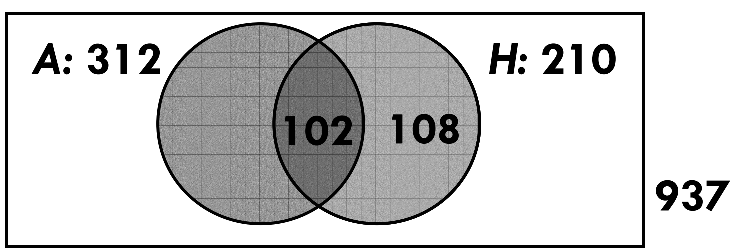


A = At least one parent with heart disease

H = Died of causes related to heart disease

$$n(A) = 312$$
$$n(\sim A) = 937 - 312 = 625$$
$$n(A \cap H) = 102$$
$$n(H \cap \sim A) = n(H) - 102 = 108$$

Solution:



A = At least one parent with heart disease

H = Died of causes related to heart disease

$$P(H \mid \sim A) = \frac{n(H \cap \sim A)}{n(\sim A)} = \frac{108}{625} = 0.173$$

Answer B

A Harder Conditional Problem:

An actuary is studying the prevalence of three health risk factors, denoted by *A, B,* and C, within a population of women. For each of the three factors, the probability is 0.1 that a woman in the population has only this risk factor (and no others). For any two of the three factors, the probability is 0.12 that she has exactly these two risk factors (but not the other).

The probability that a woman has all three risk factors, given that she has *A* and *B*, is 1/3.

25

A Harder Conditional Problem, cont.:

What is the probability that a woman has none of the three risk factors, given that she does not have risk factor *A*?

(A) 0.280
(B) 0.311
(C) 0.467
(D) 0.484
(E) 0.700

26

DeMorgan's Laws:

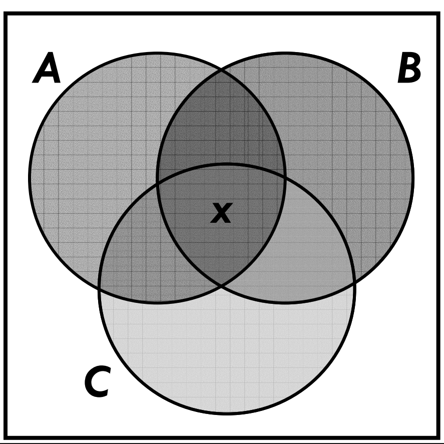
$$\sim A \cap \sim B = \sim\left(A \cup B\right)$$

$$\sim A \cup \sim B = \sim\left(A \cap B\right)$$

Harder Problem Solution:

We want to find

$$P(\sim A \cap \sim B \cap \sim C \mid \sim A) = \frac{P(\sim A \cap \sim B \cap \sim C)}{P\left(\sim A\right)}$$

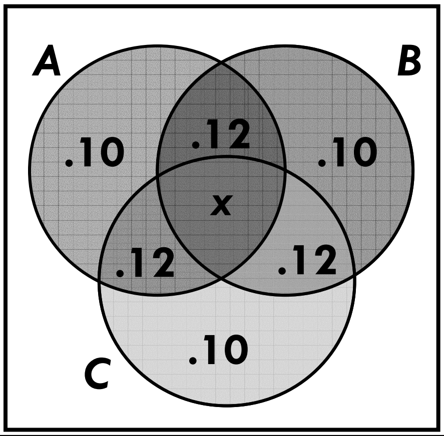


$$= \frac{P[\sim (A \cup B \cup C)]}{P\left(\sim A\right)}$$

But , $P\left(A \cap B \cap C\right) = x$ is not given.

Harder Problem Solution, cont.:

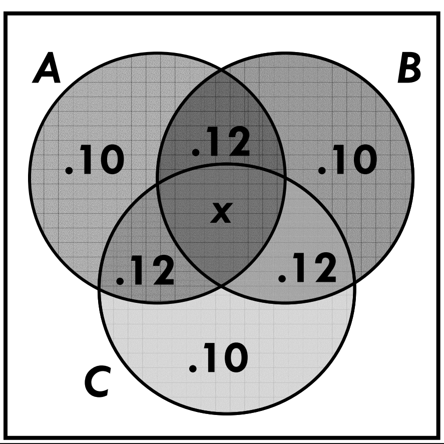
Fill in 0.12 in each of the areas representing exactly two risk factors, and fill in 0.10 in each of the areas representing exactly one risk factor.



29

Harder Problem Solution, cont.:

Probability of a woman having all three risk factors given that she has A and B is 1/3.



$$P\left(A\cap B\cap C \mid A\cap B\right) = \frac{P\left(A\cap B\cap C\right)}{P\left(A\cap B\right)} = \frac{1}{3}$$

$$P\left(A\cap B\cap C\right) = x$$

$$P\left(A\cap B\right) = x + 0.12$$

$$\frac{x}{x+0.12} = \frac{1}{3} \rightarrow x = 0.06$$

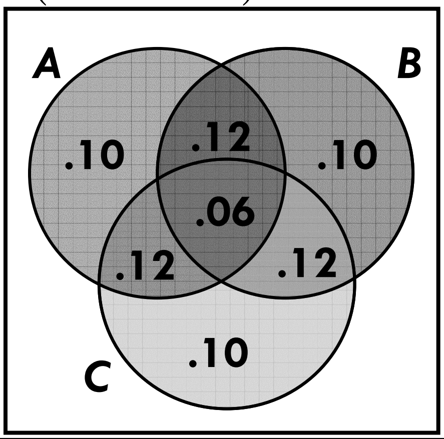
30

Harder Problem Solution, cont.:

$$P(A) = 0.06 + 0.12 + 0.12 + 0.10$$

$$= 0.40 \qquad \rightarrow \qquad P(\sim A) = 0.60$$

$$P(A \cup B \cup C) = 0.06 + 3(0.12) + 3(0.10) = 0.72$$



$$P[\sim(A \cup B \cup C)]$$

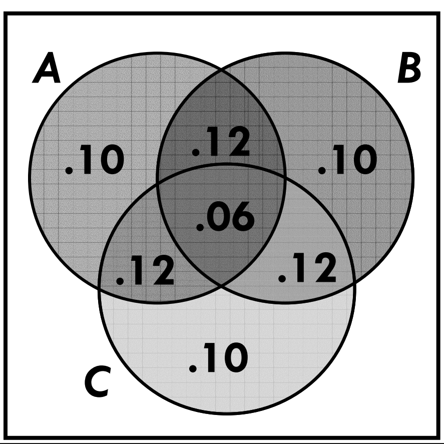
$$= 1 - 0.72$$

$$= 0.28$$

31

Harder Problem Solution, cont.:

$$P(\sim A \cap \sim B \cap \sim C \mid \sim A) = \frac{P[\sim(A \cup B \cup C)]}{P(\sim A)}$$



$$= \frac{0.28}{0.60}$$

$$= 0.467$$

Answer C

32

More Probability Rules:

- **Definition:**

  Two events A and B, are **independent** if

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

- **Multiplication Rule for Independent Events**

  If A and B, are independent,

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

33

Exercise:

An actuary studying insurance preferences makes the following conclusions:

(i) A car owner is twice as likely to purchase collision coverage as disability coverage.

(ii) The event that a car owner purchases collision coverage is independent of the event that he or she purchases disability coverage.

(iii) The probability that a car owner purchases both collision and disability coverages is 0.15 .

34

Exercise, cont.:

What is the probability that an automobile owner purchases neither collision nor disability coverage?

(A) 0.18
(B) 0.33
(C) 0.48
(D) 0.67
(E) 0.82

Solution:

Let C be collision insurance and *D* be disability insurance.

We need to find $P\left[\sim\left(C \cup D\right)\right] = 1 - P\left(C \cup D\right)$.

i) $P(C) = 2P(D)$

ii) $P\left(C \cap D\right) = P(C)P(D)$

iii) $P(C \cap D) = 0.15$

Solution, cont.:

$$0.15 = P\left(C \cap D\right) = P(C)P(D) = 2P(D)^2$$

$$P(D)^2 = 0.075 \to P(D) = \sqrt{0.075}$$

$$P(C) = 2P(D) = 2\sqrt{0.075}$$

$$\begin{aligned} P(C \cup D) &= P(C) + P(D) - P(C \cap D) \\ &= 2\sqrt{0.075} + \sqrt{0.075} - 0.15 \\ &= 0.67 \end{aligned}$$

$$P\left[\sim\left(C \cup D\right)\right] = 1 - P\left(C \cup D\right) = 1 - 0.67 = 0.33$$

Answer B

Bayes Theorem – Simplify with Trees:

A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. 1% of the population actually has the disease.

Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

(A)0.324 (B) 0.657 (C) 0.945
(D) 0.950 (E) 0.995

Solution:

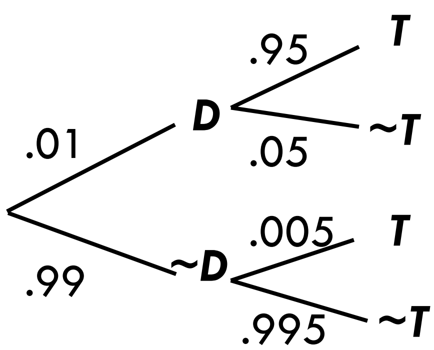
D = Person has the disease

T = Test indicates the disease

We need to find

$$P(D \mid T) = \frac{P(D \cap T)}{P(T)}$$

Solution, cont.:



$$0.0095 = P(D \cap T)$$

$$0.00495 = P(\sim D \cap T)$$

$$P(D \mid T) = \frac{P(D \cap T)}{P(T)} = \frac{0.0095}{0.0095 + 0.00495} = 0.657$$

Answer B

Probability Rules:

Law of Total Probability:

Let E be an event. If $A_1,\ A_2, \ldots A_n$ partition the sample space, then

$$P(E) = P(A_1 \cap E) + P(A_2 \cap E) + \ldots + P(A_n \cap E).$$

Theorem:

Bayes' Theorem:

Let E be an event. If $A_1,\ A_2, \ldots A_n$ partition the sample space, then

$$P(A_1 \mid E) = \frac{P(E \cap A_1)}{P(E)}$$

$$= \frac{P(A_1)P(A_1 \mid E)}{P(A_1)P(A_1 \mid E) + \cdots + P(A_n)P(A_n \mid E)}$$

Exercise:

An insurance company issues life insurance policies in three separate categories: standard, preferred, and ultra-preferred. Of the company's policyholders, 50% are standard, 40% are preferred, and 10% are ultra-preferred. Each standard policyholder has probability 0.010 of dying in the next year, each preferred policyholder has probability 0.005 of dying in the next year, and each ultra-preferred policyholder has probability 0.001 of dying in the next year.

43

Exercise, cont.:

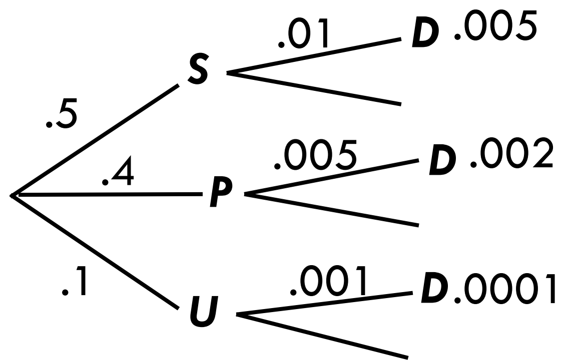
A policyholder dies in the next year.

What is the probability that the deceased policyholder was ultra preferred?

(A) 0.0001
(B) 0.0010
(C) 0.0071
(D) 0.0141
(E) 0.2817

44

Solution



$$P(U \mid D) = \frac{P(U \cap D)}{P(D)}$$

$$= \frac{.0001}{.0001 + .002 + .005} = .0141$$

Answer D

Exercise:

The probability that a randomly chosen male has a circulation problem is 0.25 . Males who have a circulation problem are twice as likely to be smokers as those who do not have a circulation problem.

What is the conditional probability that a male has a circulation problem, given that he a smoker?

(A) 1/4 (B) 1/3 (C) 2/5 (D) 1/2 (E) 2/3

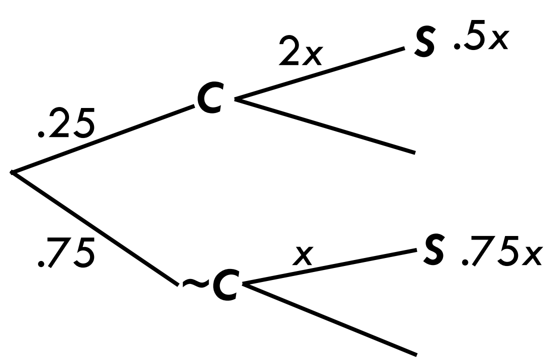
Solution:

C = Circulatory problem
S = Smoker

- We need to find $P(C \mid S)$.
- We do not know $x = P(S \mid \sim C)$.
- We do know that $2x = P(S \mid C)$ since those who have a circulation problem are twice as likely to be smokers.

47

Solution, cont.:



$$P(C \mid S) = \frac{P(C \cap S)}{P(S)} = \frac{.5x}{.5x + .75x} = \frac{.5}{1.25} = .40$$

Answer C

48

Expected Value:

- **Definition:**

  The **expected value** of X is defined by

  $$E(X) = \sum x\, p(x)$$

  The expected value is also referred to as the **mean** of the random variable X and denoted by Greek letter μ. $E(x) = \mu$.

- **A Property of Expected Value:**

  $$E(aX + b) = aE(X) + b$$

49

Variance:

- **Definition:**

  The **variance** of a random variable X is

  $$V(X) = E\left[(X - \mu)^2\right] = \sum (x - \mu)^2\, p(x)$$

- **Standard Deviation:**

  $\sigma = \sqrt{V(X).}$ Notation: $V(X) = \sigma^2$

- $V(X) = E(X^2) - E(X)^2 = E(X^2) - \mu^2$

- $V(aX + b) = a^2 V(X)$

50

Exercise:

A probability distribution of claim sizes is given in this table:

Claim Size	Probability
20	0.15
30	0.10
40	0.05
50	0.20
60	0.10
70	0.10
80	0.30

Exercise, cont.:

What percentage of the claims are within one standard deviation of the mean claim size?

(A)45%
(B) 55%
(C) 68%
(D) 85%
(E) 100%

Solution:

Claim Size	Probability	$xp(x)$	$x^2p(x)$
20	0.15	3	60
30	0.10	3	90
40	0.05	2	80
50	0.20	10	500
60	0.10	6	360
70	0.10	7	490
80	0.30	24	1920
Total	1.00	55	3500

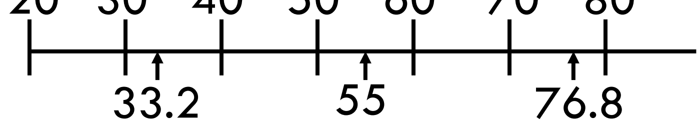
Solution, cont.:

$$E\left(X\right) = \sum x\,p\left(x\right) = 55$$

$$\sigma^2 = V\left(X\right) = E\left(X^2\right) = \mu^2 = 3500 - 55^2 = 475$$

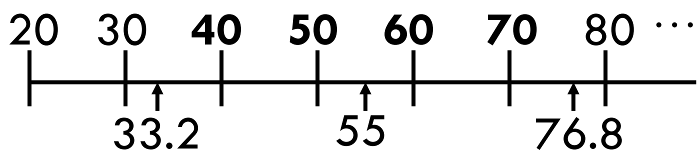
$$\sigma = \sqrt{475} = 21.8$$

A value is within one standard deviation of the mean if it is in the interval $\left[\mu - \sigma, \mu + \sigma\right]$, that is, in the interval $\left[33.2,\ 76.8\right]$.



Solution, cont.:

The values of x in this interval are 40, 50, 60, and 70.



Thus, the probability of being within one standard deviation of the mean is:

$$p(40) + p(50) + p(60) + p(70)$$
$$= .05 + .20 + .10 + .10 = .45$$

Z-Score:

Definition:

For any possible value x of a random variable,

$$z = \frac{x - \mu}{\sigma}$$

The z score measures the distance of x from $E\left(X\right) = \mu$ in standard deviation units.

Theorem:

- **Chebychev's Theorem:**

  For any random variable X, the probability that X is within k standard deviations of the mean is at least $1-\frac{1}{k^2}$.

- $P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - \frac{1}{k^2}$

57

Additional Properties of V(X):

- $V(X+Y) = V(X) + V(Y) + 2\operatorname{cov}(X,Y)$

- For X, Y independent

  $V(X+Y) = V(X) + V(Y)$

58

Exercise:

The profit for a new product is given by

$Z = 3X - Y - 5.$

X and Y are independent random variables with

$V(X) = 1$ and $V(Y) = 2.$

What is the variance of Z?

(A) 1 (B) 5 (C) 7 (D) 11 (E) 16

59

Solution:

$$
\begin{aligned}
V(Z) &= V\left(3X - Y - 5\right) = V\left(3X - Y\right) \\
&= V\left(3X + \left(-Y\right)\right) \underset{independence}{=} V\left(3X\right) + V\left(-Y\right) \\
&= 3^2 V\left(X\right) + \left(-1\right)^2 V\left(Y\right) \\
&= 9\left(1\right) + 2 = 11
\end{aligned}
$$

Answer D

Note!

Observe the wrong answer which you would obtain if you mistakenly wrote $V(3X - Y) = V(3X) - V(Y)$*.*
This is choice C, and is a common careless mistake.

60

Exercise:

A recent study indicates that the annual cost of maintaining and repairing a car in a town in Ontario averages 200 with a variance of 260.

If a tax of 20% is introduced on all items associated with the maintenance and repair of cars (*i.e.*, everything is made 20% more expensive), what will be the variance of the annual cost of maintaining and repairing a car?

(A) 208 (B) 260 (C) 270 (D) 312 (E) 374

Solution:

Let X be the random variable for the present cost, and Y=1.2X the random variable for the cost after 20% inflation. We are asked to find $V(Y)$.

$$\begin{aligned} V(Y) &= V(1.2X) \\ &= 1.2^2 V(X) \\ &= 1.44(260) \\ &= 374.4 \end{aligned}$$

Answer E

Geometric Series Review:

- A **geometric sequence** is of the form

  $a,\ ar,\ ar^2,\ ar^3,\dots,\ ar^n.$

- The sum of the series for $r \neq 1$ is given by:

  $$a + ar + ar^2 + \dots + ar^n = a\left(\frac{1-r^{n+1}}{1-r}\right)$$

  The number r is called the **ratio**.

- If $|r|<1$, we can sum the infinite geometric series:

  $$a + ar + ar^2 + \dots + ar^n + \dots = a\left(\frac{1}{1-r}\right)$$

63

Geometric Distribution:

$$P(X=k) = q^k p, \quad k = 0, 1, 2, 3, \dots$$

$$E(X) = \frac{q}{p} \qquad V(X) = \frac{q}{p^2}$$

where $X=$ the number of failures before the first success in a repeated series of independent success-failure trials with $P(\text{Success}) = p$.

64

Geometric Distribution Alternative:

Here, you are looking at the number of trials needed to get to the first success. In this formulation, you are looking at $Y = X + 1$.

$$P(Y = k) = q^{k-1}p, \quad k = 1, 2, 3, \ldots$$

$$E(Y) = \frac{1}{p} \qquad V(Y) = \frac{q}{p^2}$$

Exercise:

In modeling the number of claims filed by an individual under an automobile policy during a three-year period, an actuary makes the simplifying assumption that for all integers $n \geq 0$, $p_{n+1} = \frac{1}{5} p_n$ where p_n represents the probability that the policyholder files n claims during the period.

Under this assumption, what is the probability that a policyholder files more than one claim during the period?

(A) 0.04 (B) 0.16 (C) 0.20 (D) 0.80 (E) 0.96

Solution:

We are not given p_0. Look at the first few terms:

$$p_0,\ p_0\left(\frac{1}{5}\right),\ p_0\left(\frac{1}{5}\right)^2,\ p_0\left(\frac{1}{5}\right)^2,\ldots$$

Solution, cont.:

$$1 = p_0\left[1+\frac{1}{5}+\left(\frac{1}{5}\right)^2+\left(\frac{1}{5}\right)^3+\ldots+\left(\frac{1}{5}\right)^n+\ldots\right]$$

$$= p_0\left(\frac{1}{1-\frac{1}{5}}\right) = \frac{5}{4}p_0$$

$$\rightarrow p_0 = \frac{4}{5}$$

Solution, cont.:

$$P\left(N>1\right)=1-P\left(N\leq 1\right)=1-\left[\frac{4}{5}+\frac{4}{5}\left(\frac{1}{5}\right)\right]=.04$$

Answer A

Note!

The probability distribution has the form of a geometric distribution with $q=\frac{1}{5}$, *so it must be true that* $p_0=p=\frac{4}{5}$.

Binomial Distribution:

Given *n* independent, success-failure trials with $P(S)=p, \quad P\left(F\right)=1-p=q$

- $$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$
  $$=\binom{n}{k}p^k(q)^{n-k}, \quad k=0,1,\ldots,n$$
- $E\left(X\right)=np$

  $V\left(X\right)=np\left(1-p\right)=npq$

Notation Review:

- $n! = n(n-1)\ldots(2)1$
- $\binom{n}{r} = C(n,r) = \dfrac{P(n,r)}{r!}$

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

$$= \frac{n(n-1)\cdots(n-r+1)}{r!}$$

- $\binom{10}{2} = \dfrac{10!}{2!8!} = \dfrac{10\cdot 9}{2\cdot 1} = 45$

71

Example:

Guessing on a 10 question multiple choice quiz with choices A, B, C, D, E.

$$n = 10, \quad P(S) = .2 = p, \quad q = .8$$

$$P(X=2) = \binom{10}{2}(.2)^2(.8)^8 \approx .302$$

$$E(X) = 10(.2) = 2$$

$$V(X) = 10(.2)(.8) = 1.6$$

72

Exercise:

A study is being conducted in which the health of two independent groups of ten policyholders is being monitored over a one-year period of time. Individual participants in the study drop out before the end of the study with probability 0.2 (independently of the other participants).

What is the probability that at least 9 participants complete the study in one of the two groups, but not in both groups?

(A) .096 (B) .192 (C) .235 (D) .376 (E).469

Solution:

Denote the random variables for the number of participants completing in each group by A and B. We need

$$P\left[\left(A \geq 9 \,\&\, B < 9\right) or \left(B \geq 9 \,\&\, A < 9\right)\right]$$

$$= P\left(A \geq 9 \,\&\, B < 9\right) + P\left(B \geq 9 \,\&\, A < 9\right)$$

$$\underset{Ind}{=} P\left(A \geq 9\right)P\left(B < 9\right) + P\left(B \geq 9\right)P\left(A < 9\right)$$

The two groups are independent and have identical binomial probability distributions.

Solution, cont.:

A is binomial with n=10 independent trials and probability of completion p=0.8.

$$P\left(A\ge 9\right)=P\left(A=10\right)+P\left(A=9\right)$$

$$=.8^{10}+\binom{10}{9}.8^{9}\left(.2\right)=.376$$

$$P(A<9)=1-P\left(A\ge 9\right)=.624$$

$$P\left(B\ge 9\right)=.376\rightarrow P(B<9)=.624$$

$$P\left(A\ge 9\right)P\left(B<9\right)+P\left(B\ge 9\right)P\left(A<9\right)$$

$$=.376\left(.624\right)+.376\left(.624\right)=.469$$

Answer E

Harder Bayes Thrm./Binomial Exercise:

A hospital receives 1/5 of its flu vaccine shipments from Company X and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials.

For Company X's shipments, 10% of the vials are ineffective. For every other company, 2% of the vials are ineffective. The hospital tests 30 randomly selected vials from a shipment and finds that one vial is ineffective.

Bayes Thrm./Binomial Exercise, cont.:

What is the probability that this shipment came from Company *X*?

(A) 0.10
(B) 0.14
(C) 0.37
(D) 0.63
(E) 0.86

Solution:

X = Shipment came from company *X*
I = Exactly 1 vial out of 30 tested is ineffective

We are asked to find $P\left(X \mid I\right)$.

If the shipment is from company *X*, the number of defectives in 30 components is a binomial random variable with n=30 and p=0.1.

The probability of one defective in a batch of 30 from *X* is

$$P\left(I \mid X\right) = \binom{30}{1}\left(.1\right)\left(.9^{29}\right) = .141$$

Solution, cont.:

X = Shipment came from company X

I = Exactly 1 vial out of 30 tested is ineffective

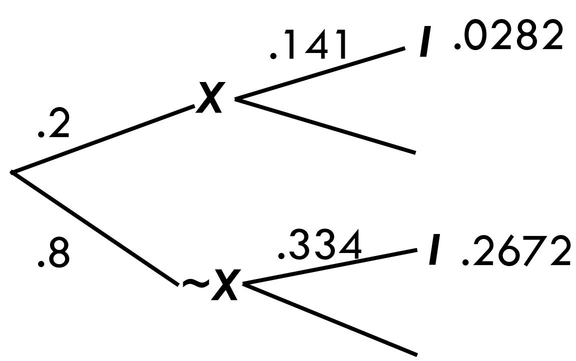
We are asked to find $P(X \mid I)$.

If the shipment isn't from company X, the number of defectives in 30 components is a binomial random variable with n=30 and p=0.02.

$$P(I \mid \sim X) = \binom{30}{1}(.02)(.98^{29}) = .334$$

79

Solution, cont.:



$$P(X \mid I) = \frac{P(X \cap I)}{P(I)} = \frac{.0282}{.0282 + .2672} = .0955$$

Answer A

80

Poisson Distribution:

X is Poisson with mean λ.

- $P(X = k) = \dfrac{e^{-\lambda}\lambda^{k}}{k!}, \quad k = 1, 2, 3, \ldots$

- $E(X) = \lambda$

  $V(X) = \lambda$

Example:

Accidents occur at an average rate of $\lambda = 2$ per month.

Let X = the number of accidents in a month.

$$P(X = 1) = \frac{e^{-2}2^{1}}{1!} \approx .271$$

$$E(X) = V(X) = 2$$

Exercise:

An actuary has discovered that policyholders are three times as likely to file two claims as to file four claims. If the number of claims filed has a Poisson distribution, what is the variance of the number of claims filed?

(A) $1/\sqrt{3}$

(B) 1

(C) $\sqrt{2}$

(D) 2

(E) 4

Solution:

$$P(X=2)=3P(X=4)$$

$$\frac{e^{-\lambda}\lambda^2}{2!}=3\left(\frac{e^{-\lambda}\lambda^4}{4!}\right)$$

$$4\lambda^2=\lambda^4 \rightarrow \lambda=2$$

Answer D

Hypergeometric Example:

A company has 20 male and 30 female employees. 5 employees are chosen at random for drug testing. What is the probability that 3 males and 2 females are chosen?

Solution:

$$\frac{\binom{20}{3}\binom{30}{2}}{\binom{50}{5}} \approx 0.234$$

Hypergeometric Probabilities:

1 A sample of size *n* is being taken from a finite population of size *N*.

2 The population has a subgroup of size $r \geq n$ that is of interest.

3 The random variable of interest is *X*, the number of members of the subgroup in the sample taken.

Hypergeometric Probabilities, cont.:

- $$P(X=k)=\frac{\binom{N-r}{n-k}\binom{r}{k}}{\binom{N}{n}}, \quad k=0,\ldots,n$$

- $$E(X)=n\left(\frac{r}{N}\right)$$

$$V(X)=n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

87

Previous Example, cont.:

X = number of males chosen in a sample of 5.

$N=50 \qquad n=5 \qquad r=20$

$$P(X=k)=\frac{\binom{30}{5-k}\binom{20}{k}}{\binom{50}{5}}$$

$$E(X) = 5\left(\frac{20}{50}\right)=2$$

$$V(X)=5\left(\frac{20}{50}\right)\left(1-\frac{20}{50}\right)\left(\frac{50-5}{50-1}\right)$$

88

Negative Binomial Distribution:

A series of independent trials has $P(S) = p$ on each trial.

Let X be the number of failures before success r.

- $P(X = k) = \binom{r+k-1}{r-1} q^k p^r, \quad k = 0, 1, 2, 3, \ldots$
- $E(X) = \dfrac{rq}{p} \qquad V(X) = \dfrac{rq}{p^2}$
- The special case with $r = 1$ is the geometric random variable.

89

Example:

Play slot machine repeatedly with probability of success on each independent play $P(S) = .05 = p$.

Find the probability of exactly 4 losses (failures) before the second win (success r=2).

90

Example, cont.:

Possible sequences:

SFFFFS
FSFFFS
FFSFFS
FFFSFS
FFFFSS

Single sequence probability : $.05^2(.95)^4$

Number of sequences: $\binom{5}{1} = 5$

Solution: $5(.05)^2(.95)^4 \approx .0818$

Exercise:

A company takes out an insurance policy to cover accidents at its manufacturing plant. The probability that one or more accidents will occur during any given month is 3/5. The number of accidents that occur in any given month is independent of the number of accidents that occur in all other months.

Exercise, cont.:

Calculate the probability that there will be at least four months in which no accidents occur before the fourth month in which at least one accident occurs.

(A) 0.01 (B) 0.12 (C) 0.23 (D) 0.29 (E) 0.41

Solution:

This is a negative binomial distribution problem.

Success S = month with at least one accident

Failure F = month with no accidents.

Note that $P(S) = p = 3/5$.

Let X be the number of months with no accidents before the fourth month with at least one accident –i.e., the number of failures before the fourth success.

X is negative binomial with $r = 4$ and $p=3/5$.

Solution, cont.:

We are asked to find

$$P\left(X \geq 4\right) = 1 - P\left(X \leq 3\right)$$

$$= 1 - \left[P\left(X=0\right) + P\left(X=1\right) + P\left(X=2\right) + P\left(X=3\right)\right]$$

$$P\left(X=0\right) = \left(\frac{3}{5}\right)^4 = 0.12960$$

$$P\left(X=1\right) = \binom{4}{1}\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right) = .20736$$

$$P\left(X=2\right) = \binom{5}{2}\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^2 = .20736$$

$$P\left(X=3\right) = \binom{6}{3}\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^3 = .16589$$

Solution, cont.:

$$P\left(X \leq 3\right) = .12960 + .20736 + .20736 + .16589$$
$$= .71021$$

$$P\left(X \geq 4\right) = 1 - P\left(X \leq 3\right)$$
$$= 1 - .71021 = .28979$$

Answer D

Definition of Continuous Distribution:

The probability density function of a random variable X is a real valued function satisfying:

1 $f(x) \geq 0$ for all x.

2 The total area bounded by the graph of $y = f(x)$ and the x axis is 1.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

3 $P\left(a \leq X \leq b\right)$ is given by the area under $y = f(x)$ between $x = a$ and $x = b$.

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$$

Continuous Distribution Properties:

- **Cumulative Distribution Function *F(x)***

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

- **Expected Value**

$$E(X) = \int_{-\infty}^{\infty} x\, f\left(x\right) dx$$

- **Expected value of a function of a continuous random variable**

$$E\left[g\left(X\right)\right] = \int_{-\infty}^{\infty} g\left(x\right) \cdot f\left(x\right) dx$$

- **Mean of *Y* = *aX* + *b***

$$E\left(aX + b\right) = a\,E\left(X\right) + b$$

Continuous Distribution Properties:

- **Variance**

  $$V(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,dx$$

- **Alternate Form of Variance Calculation**

  $$V(X) = E(X^2) - \left[E(X)\right]^2 = E(X^2) - \mu^2$$

- **Variance of *Y = aX + b***

  $$V(aX + b) = a^2 V(X)$$

99

Uniform Random Variable on [*a*, *b*]:

- $$f(x) = \begin{cases} \dfrac{1}{b - a}, a \le x \le b \\ 0, \text{ otherwise} \end{cases}$$

- $$E(X) = \frac{a+b}{2} \qquad V(X) = \frac{(b-a)^2}{12}$$

100

Exponential Distribution:

- Random variable T, parameter λ.
  T is often used to model waiting time, $\lambda = \text{rate}$.

- $f(t) = \lambda e^{-\lambda t}, \quad F(t) = 1 - e^{-\lambda t}$ for $t \geq 0$

- $E(T) = \dfrac{1}{\lambda} \qquad V(T) = \dfrac{1}{\lambda^2}$

Example:

Waiting time for next accident.
$\lambda = 2$ accidents per month on average.

$$P\left(0 \leq T \leq 1\right) = F(1) = 1 - e^{-2} \approx .865$$

$$E\left(T\right) = \frac{1}{2} \qquad V\left(T\right) = \frac{1}{4}$$

* **Exponential** waiting time
Poisson number of events *

Useful Exponential Facts:

- $\lim_{x\to\infty} x^n e^{-ax} = \lim_{x\to\infty} \frac{x^n}{e^{ax}} = 0,$ for $a > 0.$

- $\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}},$

  for $a > 0$, and n a positive integer.

Exercise:

The waiting time for the first claim from a good driver and the waiting time for the claim from a bad driver are independent and follow exponential distributions with 6 years and 3 years, respectively.

What is the probability that the first claim from a good driver will be filed within 3 years and the first claim from a bad driver will be filed within 2 years?

Exercise:

(A) $\frac{1}{18}\left(1-e^{-\frac{2}{3}}-e^{-\frac{1}{2}}+e^{-\frac{7}{6}}\right)$

(B) $\frac{1}{18}e^{-\frac{7}{6}}$

(C) $1-e^{-\frac{2}{3}}-e^{-\frac{1}{2}}+e^{-\frac{7}{6}}$

(D) $1-e^{-\frac{2}{3}}-e^{-\frac{1}{2}}+e^{-\frac{1}{3}}$

(E) $1-\frac{1}{3}e^{-\frac{2}{3}}-\frac{1}{6}e^{-\frac{1}{2}}+\frac{1}{18}e^{-\frac{7}{6}}$

Solution:

Recall, the mean of the exponential is $\mu = 1/\lambda$. Thus if you are given the mean (as in this problem), you know that $1/\mu = \lambda$.

G: Waiting time for 1st accident for good driver
B: Waiting time for 1st accident for bad driver

G: $\lambda_G = \frac{1}{6}$ $\qquad F_G(x) = \left(1-e^{-\frac{x}{6}}\right)$

B: $\lambda_B = \frac{1}{3}$ $\qquad F_B(x) = \left(1-e^{-\frac{x}{3}}\right)$

Solution, cont.:

Find $P(G \le 3 \;\; \& \;\; B \le 2)$.

Note that G and B are independent.

$$P(G \le 3 \,\&\, B \le 2) = P(G \le 3)P(B \le 2)$$
$$= F_G(3)F_B(2)$$
$$= \left(1 - e^{-\frac{3}{6}}\right)\left(1 - e^{-\frac{2}{3}}\right)$$
$$= 1 - e^{-\frac{2}{3}} - e^{-\frac{1}{2}} + e^{-\frac{7}{6}}$$

Answer C

Exercise:

The number of days that elapse between the beginning of a calendar year and the moment a high-risk driver is involved in an accident is exponentially distributed. An insurance company expects that 30% of high-risk drivers will be involved in an accident during the first 50 days of a calendar year.

What portion of high-risk drivers are expected to be involved in an accident during the first 80 days of a calendar year?

(A) 0.15 (B) 0.34 (C) 0.43 (D) 0.57 (E) 0.66

Solution:

T: time in days until the first accident for a high risk driver

To find: $P(T \leq 80) = F(80)$.

We know $F(t) = 1 - e^{-\lambda t}$, but we don't know λ.

Use the given probability for the first 50 days to find it.

Solution, cont.:

$$P(T \leq 50) = F(50)$$
$$= 1 - e^{-\lambda 50}$$
$$= 0.30$$
$$\lambda = \frac{\ln(0.7)}{-50}$$

Now we have λ and can finish the problem.

$$P(T \leq 80) = F(80) = 1 - e^{-80\lambda} = .4348$$

Answer B

Definitions:

- The **mode** of a continuous random variable is the value of x for which the density function $f(x)$ is a maximum.
- The **median** m of a continuous random variable X is defined by $F(m) = P(X \leq m) = 0.50$.
- Let X be a continuous random variable and $0 \leq p \leq 1$. The **100p^th^ percentile** of X is the number x_p defined by $F(x_p) = p$.

  Note that the 50th percentile is the median.

Exercise:

An insurance policy reimburses dental expense, X, up to a maximum benefit of 250. The probability density function for X is:

$$f(x) = \begin{cases} ce^{-0.004x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where c is constant.

Calculate the median benefit for this policy.

(A) 161 (B) 165 (C) 173 (D) 182 (E) 250

Solution:

You can see by direct examination that X must be exponential with $c = .004,$ since $.004e^{-0.004x}$ is the density function for the exponential with $\lambda = .004.$

(Some of our students integrated the density function and set the total area under the curve equal to 1, but that takes extra time.)

Solution, cont.:

Original expense X: cumulative distribution $F(x) = 1 - e^{-.004x}.$

Thus the median m for X is obtained by solving the equation

$$F(m) = 0.50 = 1 - e^{-.004m} \rightarrow 0.50 = e^{-.004m}$$

$$m = \frac{\ln(.50)}{-.004} = 173.3$$

Actual benefit capped at 250. Since 173.3 is less than 250, 50% of the benefits paid are still less than 173.3 and 50% are greater.

Answer C

Normal Random Variable:

- $\mu = E(X)$ and $\sigma^2 = V(X)$

- $f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

115

Transformation to Standard Normal:

Transform any normal random variable X with mean μ and standard deviation σ into a **standard normal random variable** Z with mean 0 and standard deviation 1.

$$Z = \frac{X-\mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$$

Then probabilities can be calculated using the standard normal probability tables for Z.

116

Normal Distribution Table:

NORMAL DISTRIBUTION TABLE

Entries represent the area under the standardized normal distribution from -∞ to z, Pr(Z<z)
The value of z to the first decimal is given in the left column. The second decimal place is given in the top rov

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]					

Standard Normal Example:

X normal, $\mu = 500, \quad \sigma = 100$

$$P\left(600 \le X \le 750\right)$$

$$= P\left(\frac{600-500}{100} \le Z \le \frac{650-500}{100}\right)$$

$$= P\left(1 \le Z \le 1.5\right)$$

$$= .9332 - .8413$$

$$= .0919$$

Central Limit Theorem:

Let $X_1, X_2, \ldots, X_n$ be independent random variables all of which have the same probability distribution and thus the same mean μ and variance σ^2. If n is large, the sum

$$S = X_1 + X_2 + \ldots + X_n$$

will be approximately normal with mean $n\mu$ and variance $n\sigma^2$.

119

Exercise:

An insurance company issues 1250 vision care insurance policies. The number of claims filed by a policyholder under a vision care insurance policy during one year is a Poisson random variable with mean 2. Assume the numbers of claims filed by distinct policyholders are independent of one another.

What is the approximate probability that there is a total of between 2450 and 2600 claims during a one-year period?

(A) 0.68 (B) 0.82 (C) 0.87 (D) 0.95 (E) 1.00

120

Solution:

X_i : number of claims on policy i, i=1, …, 1250. (Poisson)

X_i : iid with mean $\mu = 2$ and variance $\sigma^2 = 2$.

The total number of claims is

$$S = X_1 + \ldots + X_{1250}$$

By the central limit theorem, S is approximately normal with $E(S) = \mu_s = 1250(2) = 2500$

$$V(S) = \sigma_s{}^2 = 1250(2) = 2500$$

$$\sigma_S = \sqrt{2500} = 50$$

Solution, cont.:

Thus

$$P\left(2450 \le S \le 2600\right)$$

$$= P\left(\frac{2450-2500}{50} \le Z \le \frac{2600-2500}{50}\right)$$

$$= P\left(-1 \le Z \le 2\right)$$

$$= .9772 - .1581$$

$$= .8191$$

Answer B

Normal Distribution Percentiles:

- The percentiles of the standard normal can be determined from the tables. For example,

  $P(Z \leq 1.96) = .975$

  Thus the 97.5 percentile of the Z distribution is 1.96.

- Commonly used percentiles of Z:

Z	0.842	1.036	1.282	1.645	1.960	2.326	2.576
$P(Z<z)$	0.800	0.850	0.900	0.950	0.975	0.990	0.995

Example:

X: normal random variable with mean μ and and standard deviation σ.

Find x_p the 100p^th percentile of X using the 100p^th percentile of Z.

$$z_p = \frac{x_p - \mu}{\sigma} \to x_p = \mu + z_p \sigma$$

For example, if X is a standard test score random variable with mean $\mu = 500$ and standard deviation $\sigma = 100$ then the 99^th percentile of X is

$$x_{.99} = \mu + z_{.99}\sigma = 500 + 2.326(100) = 732.6$$

Exercise:

A charity receives 2025 contributions. Contributions are assumed to be independent and identically distributed with mean 3125 and standard deviation 250.

Calculate the approximate 90th percentile for the distribution of the total contributions received.

(A) 6,328,000 **(B) 6,338,000**
(C) 6,343,000 **(D) 6,784,000**
(E) 6,977,000

Solution:

X_i : number of contributions i, i=1, ..., 2025.

X_i : iid with mean $\mu = 3125$ and variance
$\sigma^2 = (250)^2$.

The total contribution is
$S = X_1 + \ldots + X_{2025}$

By the central limit theorem, S is approximately normal with

$$E(S) = \mu_s = 3125(2025) = 6,328,125$$

$$V(S) = \sigma_s{}^2 = 250(2025) = 126,562,500$$

$$\sigma_S = \sqrt{126,562,500} = 11,250$$

Solution, cont.:

Since $z_{.90} = 1.282$, the 90th percentile of S is

$$s_{.90} = 6{,}328{,}125 + 1.282(11{,}250)$$
$$= 6{,}342{,}547.5$$

Answer C

Theorem:

If $X_1, X_2, \ldots, X_n$ are independent normal random variables with respective means $\mu_1, \mu_2, \ldots, \mu_n$ and respective variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, then $X_1 + X_2 + \ldots + X_n$ is normal with mean $\mu_1 + \mu_2 + \ldots + \mu_n$ and variance $\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$.

Note that this shows that you don't need large n (as required by the Central Limit Theorem) to have a normal sum.

Corollary:

Let $X_1, X_2, \ldots, X_n$ be independent normal random variables all of which have the same probability distribution and thus the same mean μ and variance σ^2.

For any *n*, the sum $S = X_1 + X_2 + \ldots + X_n$ will be normal with mean $n\mu$ and variance $n\sigma^2$.

129

Corollary applied to sample mean $\overline{X}$:

Let $X_1, X_2, \ldots, X_n$ be iid normal random variables with mean μ and variance σ^2. The sample mean is defined to be

$$\overline{X} = \frac{S}{n} = \frac{X_1 + \ldots + X_n}{n}$$

For any *n*, the sample mean $\overline{X}$ will be normal with mean μ and variance $\dfrac{\sigma^2}{n}$.

130

Exercise:

Claims filed under auto insurance policies follow a normal distribution with mean 19,400 and standard deviation 5,000.

What is the probability that the average of 25 randomly selected claims exceeds 20,000?

(A) 0.01 (B) 0.15 (C) 0.27 (D) 0.33 (E) 0.45

Solution:

X_i :claim amount on policy i, i=1, …, 25.

X_i :iid with $\mu = 19,400$ and variance $\sigma^2 = 5000^2$.

The average of 25 randomly selected claims is

$$\overline{X} = \frac{S}{25} = \frac{X_1 + ... + X_{25}}{25}$$

$$E(\overline{X}) = \mu = 19,400$$

$$V(\overline{X}) = \frac{\sigma^2}{25} = \frac{5000^2}{25} = 1000^2$$

$$\sigma_{\overline{X}} = \sqrt{1000^2} = 1000$$

Solution, cont.:

$$P\left(20{,}000 < \overline{X}\right) = P\left(\frac{20{,}000 - 19{,}400}{1{,}000} < Z\right)$$
$$= P\left(.6 < Z\right)$$
$$= .2743$$

Answer C

Exercise:

A company manufactures a brand of light bulb with a lifetime in months that is normally distributed with mean 3 and variance 1. A consumer buys a number of these bulbs with the intention of replacing them successively as they burn out. The light bulbs have independent lifetimes.

Exercise:

What is the smallest number of bulbs to be purchased so that the succession of light bulbs produces light for at least 40 months with probability at least 0.9772?

(A) 14 (B) 16 (C) 20 (D) 40 (E) 55

135

Solution:

X_i : lifetime of light bulb i, i=1, …, n.

X_i : iid with $\mu = 3$ and variance $\sigma^2 = 1$.

Total lifetime of the succession of n bulbs is

$$S = X_1 + X_2 + \ldots + X_n$$

$$E(S) = \mu_S = 3n$$

$$V(S) = \sigma_S^2 = n(1) = n$$

$$\sigma_S = \sqrt{n}$$

The succession of light bulbs produces light for at least 40 months with probability at least 0.9772 .

136

Solution, cont.:

$$.9772 = P\left(S \geq 40\right) = P\left(\frac{S-3n}{\sqrt{n}} \geq \frac{40-3n}{\sqrt{n}}\right)$$

$$= P\left(Z \geq \frac{40-3n}{\sqrt{n}}\right)$$

Z tables: $P\left(Z \geq -2\right) = .9772.$

$$\frac{40-3n}{\sqrt{n}} = -2 \rightarrow 3n - 2\sqrt{n} - 40 = 0$$

Make the substitution $x = \sqrt{n}$.

$$3x^2 - 2x - 40 = 0 \rightarrow x = \sqrt{n} = 4 \rightarrow n = 16$$

Answer B

137

Definition:

The **pure premium** for an insurance is the expected value of the amount paid on the insurance. The amount paid is usually a function of a random variable *g(X)*, so to find pure premiums we use the theorem

$$E\left[g\left(x\right)\right] = \int_{-\infty}^{\infty} g\left(x\right) f\left(x\right) dx$$

138

Insurance with a cap or policy limit:

An insurance policy reimburses a loss up to a benefit limit of 10 . The policyholder's loss, Y, follows a distribution with density function:

$$f(y) = \begin{cases} \dfrac{2}{y^3} & y > 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the expected value of the benefit paid under the insurance policy?

(A) 1.0 (B) 1.3 (C) 1.8 (D) 1.9 (E) 2.0

139

Solution:

Let B=the random variable for the benefit paid.

$$B = \begin{cases} y, & 1 < y < 10 \\ 10, & 10 \le y \end{cases}$$

$$E(B) = \int_1^{10} y\left(\frac{2}{y^3}\right)dy + \int_{10}^{\infty} 10\left(\frac{2}{y^3}\right)dt$$

$$= -2y^{-1}\Big|_1^{10} - 10y^{-2}\Big|_{10}^{\infty}$$

$$= 2\left[1 - \frac{1}{10}\right] - 10\left[0 - \frac{1}{100}\right]$$

$$= 1.9$$

Answer D

140

Insurance with a deductible:

The owner of an automobile insures it against damage by purchasing an insurance policy with a deductible of 250 . In the event that the automobile is damaged, repair costs can be modeled by a uniform random variable on the interval (0, 1500) .

Determine the standard deviation of the insurance payment in the event that the automobile is damaged.

(A) 361 (B) 403 (C) 433 (D) 464 (E) 521

Solution:

X: repair cost; *Y*: amount paid by insurance.

Find the standard deviation $\sigma_Y = \sqrt{V(Y)}$.

$$Y = \begin{cases} 0, & 0 < x \leq 250 \\ x - 250, & 250 < x \end{cases}$$

Density function of *X* is $f(x) = 1/1500$ on the interval (0, 1500).

Solution, cont.:

$$E(Y)$$

$$= \int_0^{250} 0\left(\frac{1}{1500}\right)dx + \int_{250}^{1500} (x-250)\left(\frac{1}{1500}\right)dx$$

$$= \left.\frac{(x-250)^2}{3000}\right|_{250}^{1500}$$

$$= 520.833$$

143

Solution, cont.:

$$E(Y^2)$$

$$= \int_0^{250} 0\left(\frac{1}{1500}\right)dx + \int_{250}^{1500} (x-250)^2\left(\frac{1}{1500}\right)dx$$

$$= \left.\frac{(x-250)^3}{4500}\right|_{250}^{1500}$$

$$= 434{,}027.778$$

144

Solution, cont.:

$$V(Y) = E(Y^2) - E(Y)^2$$

$$= 434{,}027.778 - 520.833^2$$

$$= 162{,}760.764$$

$$\sigma_Y = \sqrt{V(Y)}$$

$$= \sqrt{162{,}760.76}$$

$$= 403.436$$

Answer B

Exercise:

A manufacturer's annual losses follow a distribution with density function

$$f(x) = \begin{cases} \dfrac{2.5(0.6)^{2.5}}{x^{3.5}} & \text{for } x > 0.6 \\ 0 & \text{otherwise} \end{cases}$$

To cover its losses, the manufacturer purchases an insurance policy with an annual deductible of 2.

What is the mean of the manufacturer's annual losses not paid by the insurance policy?

(A) 0.84 (B) 0.88 (C) 0.93 (D) 0.95 (E) 1.00

Solution:

X: actual cost; Y: part of loss not paid by policy. Find $E(Y)$.

Since there is a deductible of 2,

$$Y = \begin{cases} x, & .6 < x < 2 \\ 2, & x \geq 2 \end{cases}$$

147

Solution, cont.:

$$E(Y) = \int_{.6}^{2} x\left(\frac{2.5(0.6)^{2.5}}{x^{3.5}}\right)dx + \int_{2}^{\infty} 2\left(\frac{2.5(0.6)^{2.5}}{x^{3.5}}\right)dx$$

$$= \left.\frac{2.5(0.6)^{2.5}x^{-1.5}}{-1.5}\right|_{.6}^{2} + \left.\frac{5(0.6)^{2.5}x^{-2.5}}{-2.5}\right|_{2}^{\infty}$$

$$= .83568 + .09859$$

$$= .93427$$

Answer C

148

Exercise:

An insurance policy is written to cover a loss, X, where X has a uniform distribution on [0, 1000].

At what level must a deductible be set in order for the expected payment to be 25% of what it would be with no deductible?

(A) 250 (B) 375 (C) 500 (D) 625 (E) 750

Solution:

d: unknown deductible;

Y: amount paid by insurance.

$$Y = \begin{cases} 0, & x < d \\ x - d, & x \geq d \end{cases}$$

Density function for X: $f(x) = \dfrac{1}{1000}$, $0 \leq x \leq 1000$

$$E(Y) = \int_0^d 0\left(\frac{1}{1000}\right)dx + \int_d^{1000}(x-d)\left(\frac{1}{1000}\right)dx$$

$$= \left.\frac{(x-d)^2}{2000}\right|_d^{1000} = \frac{(1000-d)^2}{2000}$$

Solution, cont.:

Find d such that $E(Y) = .25E(X)$.

For the uniform X on [0, 1000], $E(X) = 500$ and $.25E(X) = 125$.

$$E(Y) = .25E(X)$$

$$\frac{(1000-d)^2}{2000} = 125$$

$$(1000-d)^2 = 250{,}000 \to \quad d = 500$$

Answer C 151

Finding the Density Function for Y=g(x):

Example:

Cost, X, is exponential with $\lambda = .01$. After inflation of 5%, the new cost is $Y = 1.05X$.
Find $F_Y(y)$.

Note that $F_X(x) = 1 - e^{-.01x}$.

$$F_Y(y) = P(Y \le y) = P(1.05X \le Y)$$

$$= P\left(X \le \frac{Y}{1.05}\right) = F_X\left(\frac{Y}{1.05}\right)$$

$$= 1 - e^{-.01\frac{y}{1.05}}$$

152

Example, cont.:

Y is exponential with $\lambda = \dfrac{.01}{1.05}$.

Density function: $f_Y(y) = F_Y'(y)$

Useful notation: $S(x) = P(X > x) = 1 - F(x)$

153

Density Function When Inverse Exists:

Case 1. $g(x)$ is strictly increasing on the sample space for X.

Let $h(y)$ be the inverse function of $g(x)$. The function $h(y)$ will also be strictly increasing. In this case, we can find $F_Y(y)$ as follows:

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(g(X) \le y) \\ &= P\left[h\left(g(X)\right) \le h(y)\right] \\ &= P\left(X \le h(y)\right) \\ &= F_X\left(h(y)\right) \end{aligned}$$

154

Density Function When Inverse Exists:

Case 2. **$g(x)$ is strictly decreasing on the sample space for X.**

Let $h(y)$ be the inverse function of $g(x)$. The function $h(y)$ will also be strictly decreasing. In this case, we can find $F_Y(y)$ as follows:

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(g(X) \le y) \\ &= P\left[h\left(g\left(X\right)\right) \ge h(y)\right] \\ &= P\left(X \ge h(y)\right) \\ &= S_X\left(h(y)\right) \end{aligned}$$

155

Density Function When Inverse Exists:

We can find the density function $f_Y(y)$ by differentiating $F_Y(y)$.

The final result can be written in the same way for both cases:

$$f_Y(y) = f_X\left(h(y)\right)\left|h'(y)\right|$$

156

Exercise:

The time, T, that a manufacturing system is out of operation has cumulative distribution function

$$F(t) = \begin{cases} 1 - \left(\dfrac{2}{t}\right)^2 & \text{for } t > 2 \\ 0 & \text{otherwise} \end{cases}$$

The resulting cost to the company is $Y = T^2$. Determine the density function of Y, for $y > 4$.

(A) $\dfrac{4}{y^2}$ **(B)** $\dfrac{8}{y^{3/2}}$ **(C)** $\dfrac{8}{y^3}$ **(D)** $\dfrac{16}{y}$ **(E)** $\dfrac{1024}{y^5}$

Solution:

First find the cumulative distribution function for Y:

$$F_Y\left(y\right) = P\left(Y \le y\right) = P\left(T^2 \le y\right) = P\left(T \le \sqrt{y}\right)$$

$$= F_T\left(\sqrt{y}\right) = 1 - \left(\frac{2}{\sqrt{y}}\right)^2 = 1 - \frac{4}{y}$$

Then the density function for Y is:

$$f_Y(y) = \frac{d}{dy} F_Y\left(y\right) = \frac{d}{dy}\left(1 - \frac{4}{y}\right) = \frac{4}{y^2}$$

Answer A

Exercise:

An investment account earns an annual interest rate R that follows a uniform distribution on the interval (0.04, 0.08). The value of a 10,000 initial investment in this account after one year is given by $V = 10{,}000\,e^{R}$.

Determine the cumulative distribution function, $F(v)$, of V for values of v that satisfy $0<F(v)<1$.

Exercise, cont.:

(A) $\dfrac{10{,}000e^{v/10{,}000} - 10{,}408}{425}$

(B) $25e^{v/10{,}000} - 0.04$

(C) $\dfrac{v - 10{,}408}{10{,}833 - 10{,}408}$

(D) $\dfrac{25}{v}$

(E) $25\left[\ln\left(\dfrac{v}{10{,}000}\right) - .04\right]$

Solution:

Uniform distribution fact to use here: $F(x) = \dfrac{x-a}{b-a}$.

R is uniform on (0.04, 0.08)

$$F_R\left(r\right) = \frac{r-.04}{.04},\ \text{for } 0.04 \le r \le 0.08.$$

Find the cumulative distribution function for V.

161

Solution, cont.:

$$F\left(v\right) = P\left(V \le v\right) = P\left(10{,}000e^R \le v\right)$$

$$= P\left(R \le \ln\left(\frac{v}{10{,}000}\right)\right)$$

$$= F_R\left(\ln\left(\frac{v}{10{,}000}\right)\right) = \frac{\ln\left(\frac{v}{10{,}000}\right) - .04}{.04}$$

$$= 25\left[\ln\left(\frac{v}{10{,}000}\right) - .04\right].$$

Answer E

162

Independent Random Variable Results:

General results for the minimum or maximum of two independent random variables:

Recall that the survival function of a random variable X is $S_X(t) = P(X > t) = 1 - F_X(t)$.

Recall that for X exponential we have $F(x) = 1 - e^{-\lambda x}$ and $S(x) = e^{-\lambda x}$.

163

Independent Random Variable Results:

X and Y independent random variables.

Find survival function for Min=min(X, Y):

$$\begin{aligned} S_{Min}(t) &= P(\min(X,Y) > t) \\ &= P(X > t \,\&\, Y > t) \underset{independence}{=} P(X > t) \cdot P(Y > t) \\ &= S_X(t)\, S_Y(t) \end{aligned}$$

164

Independent Random Variable Results:

<u>X and Y independent random variables.</u>

Find cumulative distribution for Max=max(X, Y):

$$
\begin{aligned}
F_{Max}(t) &= P(\max(X, Y) \le t) \\
&= P(X \le t \,\&\, Y \le t) \underset{independence}{=} P(X \le t) \cdot P(Y \le t) \\
&= F_X(t) F_Y(t)
\end{aligned}
$$

Exponential Random Variable Results:

Minimum of independent exponential random variables:

X and Y with parameters β and λ.

$$S_{Min}(t) = S_X(t) S_Y(t) = e^{-\beta t} e^{-\lambda t} = e^{-(\beta+\lambda)t}$$

Min=min(X, Y) is exponential with parameter $\beta + \lambda$.

Exercise:

In a small metropolitan area, annual losses due to storm, fire, and theft are assumed to be independent, exponentially distributed random variables with respective means 1.0, 1.5, and 2.4 .

Determine the probability that the maximum of these losses exceeds 3.

(A) 0.002 **(B) 0.050**
(C) 0.159 **(D) 0.287**
(E) 0.414

167

Solution:

X_1, X_2, X_3 : losses due to storm, fire, and theft, respectively.

Find $P\left[Max > 3\right]$, where $Max = \max\left(X_1, X_2, X_3\right)$:

$$F_{Max}(t) = F_{X_1}\left(t\right)F_{X_2}\left(t\right)F_{X_3}\left(t\right)$$
$$= \left(1 - e^{-x}\right)\left(1 - e^{-x/1.5}\right)\left(1 - e^{-x/2.4}\right)$$
$$P(Max \le 3) = F_{Max}(3)$$
$$= \left(1 - e^{-3}\right)\left(1 - e^{-3/1.5}\right)\left(1 - e^{-3/2.4}\right) = .586$$
$$P\left[Max > 3\right] = 1 - .586 = .414$$

Answer E

168

Moments of a Random Variable:

- **Definition:**
  The n^{th} moment of X is $E\left(X^n\right)$.

- **Moment Generating Function:**
  Let X be a discrete random variable. The moment generating function $M_X\left(t\right)$ is defined by

$$M_X(t) = E(e^{tX}) = \sum e^{tx} p(x)$$

169

Finding the nth moment:

Finding the nth moment using the moment generating function:

$$M_X^{(n)}(0) = \sum x^n p(x) = E(X^n)$$

170

Discrete Random Variable Example:

x	0	1	2
p(x)	.5	.3	.2
e^{tx}	$e^{0t}=1$	e^{1t}	e^{2t}

$$M_X(t) = 1(.5) + e^t(.3) + e^{2t}(.2)$$

$$M_X'(t) = 0 + e^t(.3) + e^{2t}(.2)(2)$$

$$M_X'(0) = 0 + 1(.3) + 1(.2)(2) = E(X)$$

MGF Useful Properties:

$$M_{aX+b}(t) = e^{tb}M_X(at)$$

If a random variable X has the moment generating function of a known distribution, then X has that distribution.

For X and Y independent, $M_{X+Y}(t) = M_X(t)M_Y(t)$.

Exercise:

Let X_1, X_2, X_3 be a random sample from a discrete distribution with probability function

$$p(x) = \begin{cases} \dfrac{1}{3} & \text{for } x = 0 \\ \dfrac{2}{3} & \text{for } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the moment generating function, $M(t)$ of $Y = X_1 X_2 X_3$.

Exercise, cont.:

(A) $\dfrac{19}{27} + \dfrac{8}{27}e^t$

(B) $1 + 2e^t$

(C) $\left(\dfrac{1}{3} + \dfrac{2}{3}e^t\right)^3$

(D) $\dfrac{1}{27} + \dfrac{8}{27}e^{3t}$

(E) $\dfrac{1}{3} + \dfrac{2}{3}e^{3t}$

Solution:

Since each X_i can be only 0 or 1, the product $Y = X_1, X_2, X_3$ can be only 0 or 1. In addition, Y is 1 if and only if all of the X_i are 1. Thus

$$P(Y=1) = \left(\frac{2}{3}\right)^3 = \frac{8}{27}$$

$$P(Y=0) = 1 - P(Y=1) = 1 - \left(\frac{2}{3}\right)^3 = \frac{19}{27}$$

$$M_Y(t) = E\left(e^{Yt}\right) = \frac{19}{27}e^{0t} + \frac{8}{27}e^{1t} = \frac{19}{27} + \frac{8}{27}e^{t}$$

Answer A 175

Exercise:

An actuary determines that the claim size for a certain class of accidents is a random variable, X, with moment generating function

$$M_X(t) = \frac{1}{(1-2500t)^4}$$

Determine the standard deviation of the claim size for this class of accidents.

(A)1,340 **(B) 5,000** **(C) 8,660**
(D) 10,000 **(E) 11,180** 176

Solution:

Use the derivatives of the moment generating function to find the first two moments and thus obtain $V(X)=E(X^2)-E(X)^2$.

$$M_X(t)=(1-2500t)^{-4}$$

$$M_X{}'(t)=-4(1-2500t)^{-5}(-2500)$$

$$=10{,}000(1-2500t)^{-5}$$

$$M_X{}''(t)=-50{,}000(1-2500t)^{-6}(-2500)$$

$$=125{,}000{,}000(1-2500t)^{-5}$$

177

Solution, cont.:

$$M_X{}'(0)=10{,}000$$

$$M_X{}''(0)=125{,}000{,}000$$

$$V(X)=E(X^2)-E(X)^2$$

$$=125{,}000{,}000-10{,}000^2$$

$$=25{,}000{,}000$$

$$\sigma_X=\sqrt{V(X)}=\sqrt{25{,}000{,}000}=5{,}000$$

Answer B

178

Exercise:

A company insures homes in three cities, *J*, *K*, and *L*. Since sufficient distance separates the cities, it is reasonable to assume that the losses occurring in these cities are independent.

The moment generating functions for the loss distributions of the cities are:

$$M_J(t) = (1-2t)^{-3}$$

$$M_K(t) = (1-2t)^{-2.5}$$

$$M_L(t) = (1-2t)^{-4.5}$$

179

Exercise, cont.:

Let X represent the combined losses from the three cities. Calculate $E\left(X^3\right)$.

(A)1,320 **(B) 2,082** **(C) 5,760**
(D) 8,000 **(E) 10,560**

180

Solution:

Recall that $E\left(X^3\right)=M_X'''(0)$. First find $M_X(t)$. Note that $X=J+K+L$ where summands are independent. Thus

$$\begin{aligned} M_X(t) &= M_{J+K+L}(t) \\ &= M_J(t)M_K(t)M_L(t) \\ &= (1-2t)^{-3}(1-2t)^{-2.5}(1-2t)^{-4.5} \\ &= (1-2t)^{-10} \end{aligned}$$

Solution, cont.:

$$M_X'(t)=-10(1-2t)^{-11}(-2)=20(1-2t)^{-11}$$

$$M_X''(t)=-220(1-2t)^{-12}(-2)=440(1-2t)^{-12}$$

$$\begin{aligned} M_X'''(t) &= -12(440)(1-2t)^{-13}(-2) \\ &= 10,560(1-2t)^{-13} \end{aligned}$$

$$E\left(X^3\right)=M_X'''(0)=10,560$$

Answer E

Discrete Joint Probability Function:

- **Definition:**
  Let X and Y be discrete random variables. The **joint probability function** for X and Y is the function
  $$p(x, y) = P(X = x, Y = y).$$
  Note that: $\sum_x \sum_y p(x, y) = 1$
- **Definition:**
  The **marginal probability functions** of X and Y are defined by
  $$p_X(x) = \sum_y p(x, y) \qquad p_Y(y) = \sum_x p(x, y)$$

183

Exercise:

A car dealership sells 0, 1, or 2 luxury cars on any day. When selling a car, the dealer also tries to persuade the customer to buy an extended warranty for the car.

Let X denote the number of luxury cars sold in a given day, and let Y denote the number of extended warranties sold.

184

Exercise, cont.:

$$P(X=0,\ Y=0)=1/6$$
$$P(X=1,\ Y=0)=1/12$$
$$P(X=1,\ Y=1)=1/6$$
$$P(X=2,\ Y=0)=1/12$$
$$P(X=2,\ Y=1)=1/3$$
$$P(X=2,\ Y=2)=1/6$$

What is the variance of X?

(A) 0.47 (B) 0.58 (C) 0.83 (D) 1.42 (E) 2.58

Solution:

First put the given information into a bivariate table and fill in the marginal probabilities for X.

$Y \backslash X$	0	1	2
0	1/6	1/12	1/12
1	0	1/6	1/3
2	0	0	1/6
$p_X(x)$	1/6=2/12	3/12	7/12

Solution, cont.:

$$E(X) = 0\left(\frac{2}{12}\right) + 1\left(\frac{3}{12}\right) + 2\left(\frac{7}{12}\right) = \frac{17}{12}$$

$$E(X^2) = 0^2\left(\frac{2}{12}\right) + 1^2\left(\frac{3}{12}\right) + 2^2\left(\frac{7}{12}\right) = \frac{31}{12}$$

$$V(X) = \frac{31}{12} - \left(\frac{17}{12}\right)^2 = .576$$

Answer B

187

Definition:

The **joint probability density function** for two continuous random variables X and Y is a continuous, real valued function $f(x,y)$ satisfying:

i) $f(x,y) \geq 0$ for all x,y.

188

Definition:

The **joint probability density function** for two continuous random variables X and Y is a continuous, real valued function $f(x,y)$ satisfying:

ii) The total volume bounded by the graph of $z = f(x,y)$ and the x-y plane is 1.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

189

Definition:

The **joint probability density function** for two continuous random variables X and Y is a continuous, real valued function $f(x,y)$ satisfying:

iii) $P(a \le X \le b,\ c \le Y \le d)$ is given by the volume between the surface $z = f(x,y)$ and the region in the x-y plane bounded by $x = a$, $x = b$, $y = c$ and $y = d$.

$$P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f(x,y)\,dy\,dx$$

190

Definition:

Let *f(x,y)* be the joint density function for the continuous random variables *X* and *Y*. The **marginal distribution functions** of *X* and *Y* are defined by:

$$f_X\left(x\right)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$f_Y\left(y\right)=\int_{-\infty}^{\infty} f(x,y)dx$$

Exercise:

A device contains two components. The device fails if either component fails. The joint density function of the lifetimes of the components, measured in hours, is $f(s,t)$, where $0 < s < 1$ and $0 < t < 1$.

What is the probability that the device fails during the first half hour of operation?

Exercise, cont.:

(A) $\int_0^{0.5}\int_0^{0.5} f(s,t)\,ds\,dt$

(B) $\int_0^{1}\int_0^{0.5} f(s,t)\,ds\,dt$

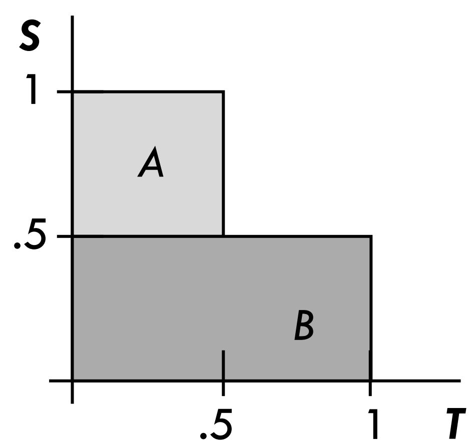
(C) $\int_{0.5}^{1}\int_{0.5}^{1} f(s,t)\,ds\,dt$

(D) $\int_0^{0.5}\int_0^{1} f(s,t)\,ds\,dt + \int_0^{1}\int_0^{0.5} f(s,t)\,ds\,dt$

(E) $\int_0^{0.5}\int_{0.5}^{1} f(s,t)\,ds\,dt + \int_0^{1}\int_0^{0.5} f(s,t)\,ds\,dt$

Solution:

The device fails if either $S < 1/2$ or $T < 1/2$.



Solution, cont.:

$$P(S<1/2 \text{ or } Y<1/2)$$
$$=\iint_A f(s,t)\,ds\,dt+\iint_B f(s,t)\,ds\,dt$$
$$=\int_0^{0.5}\int_{0.5}^{1} f(s,t)\,ds\,dt+\int_0^{1}\int_0^{0.5} f(s,t)\,ds\,dt$$

Answer E

Exercise:

The future lifetimes (in months) of two components of a machine have the following joint density function:

$$f(x,y)=\begin{cases}\dfrac{6(50-x-y)}{125{,}000}, & 0<x<50-y<50\\ 0, & \text{otherwise}\end{cases}$$

What is the probability that both components are still functioning 20 months from now?

Exercise, cont.:

(A) $$\frac{6}{125{,}000}\int_0^{20}\int_0^{20}(50-x-y)\,dy\,dx$$

(B) $$\frac{6}{125{,}000}\int_{20}^{30}\int_{20}^{50-x}(50-x-y)\,dy\,dx$$

(C) $$\frac{6}{125{,}000}\int_{20}^{30}\int_{20}^{50-x-y}(50-x-y)\,dy\,dx$$

(D) $$\frac{6}{125{,}000}\int_{20}^{50}\int_{20}^{50-x}(50-x-y)\,dy\,dx$$

(E) $$\frac{6}{125{,}000}\int_{20}^{50}\int_{20}^{50-x-y}(50-x-y)\,dy\,dx$$

Solution:

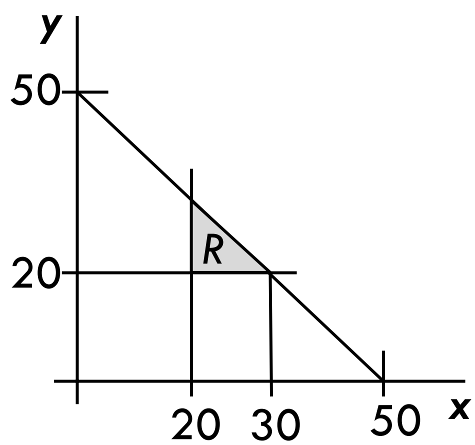
Upper limits of integration in choices C and *E* are clearly incorrect.

We need $P(X \geq 20 \,\&\, Y \geq 20)$ from *A, B* or *D*.

Density function is non-zero only in the first quadrant triangle bounded above by the line $x + y = 50$ or $y = 50 - x$.

Solution, cont.:

In the diagram, below, we show the triangle and the region R where both components are still functioning after 20 months.



Solution, cont.:

$$P\left(X \geq 20 \,\&\, Y \geq 20\right)$$
$$= \iint_R f\left(x, y\right) dy\, dx$$
$$= \frac{6}{125{,}000} \int_{20}^{30} \int_{20}^{50-x} \left(50 - x - y\right) dy\, dx$$

Answer B

Exercise:

A device runs until either of two components fails, at which point the device stops running. The joint density function of the lifetimes of the two components, both measured in hours, is

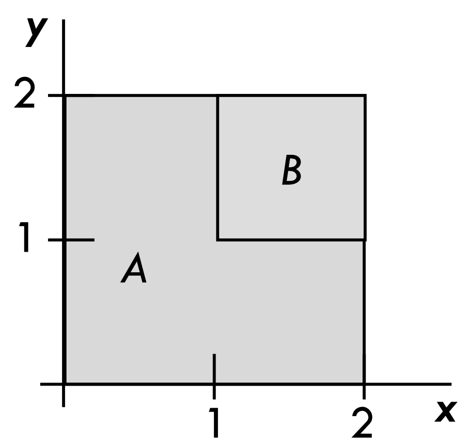
$$f(x,y) = \frac{x+y}{8} \text{ for } 0 < x < 2 \text{ and } 0 < y < 2$$

What is the probability that the device fails during its first hour of operation?

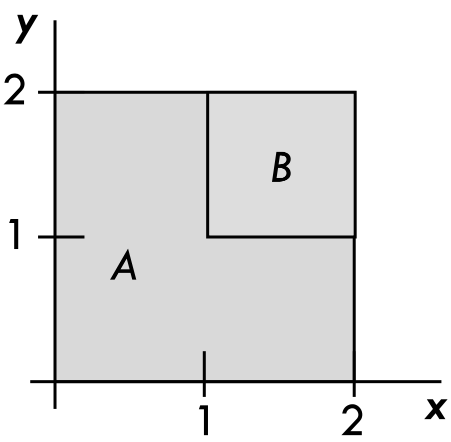
(A) .125 (B) .141 (C) .391 (D) .625 (E) .875

Solution:

The device fails if either $X < 1$ or $Y < 1$.
The set of pairs (x,y) for which this occurs is shown in the shaded region in the diagram below.



Solution, cont.:



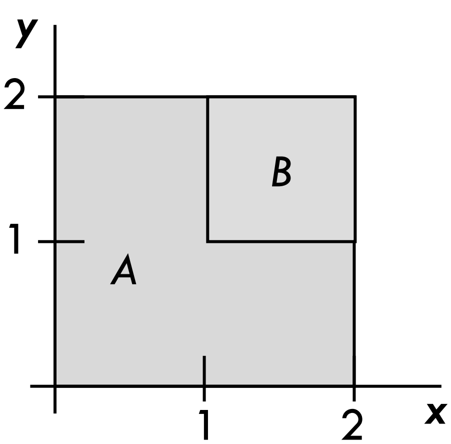
For the shaded region A,

$$P\left(X<1 \text{ or } Y<1\right)$$

$$=\iint_A f\left(x,y\right)dx\,dy$$

203

Solution, cont.:



Integrate over the unshaded rectangle B, to get the complementary probability.

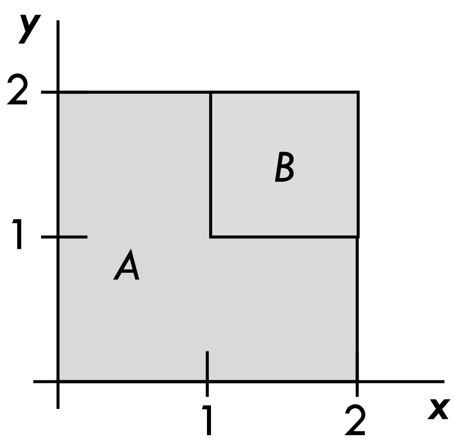
$$P\left(X<1 \text{ or } Y<1\right)$$

$$=1-\iint_B f\left(x,y\right)dx\,dy$$

$$=1-\int_1^2\int_1^2\left(\frac{x+y}{8}\right)dx\,dy$$

204

Solution, cont.:



$$\int_1^2 \int_1^2 \left(\frac{x+y}{8} \right) dx\, dy$$

$$= \frac{1}{8} \int_1^2 \left(\frac{x^2}{2} + xy \right) \Bigg|_1^2 dy$$

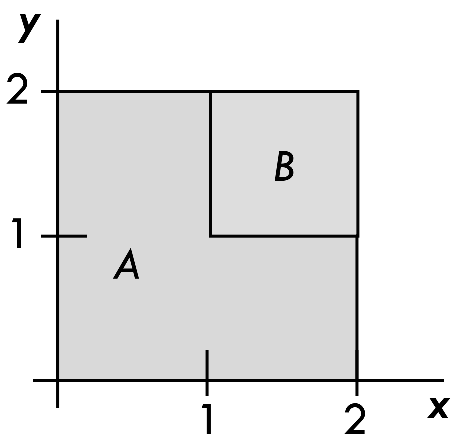
$$= \frac{1}{8} \int_1^2 (1.5 + y)\, dy$$

$$= \frac{1}{8} \left(1.5y + \frac{y^2}{2} \right) \Bigg|_1^2 = \frac{3}{8}$$

$$= .375$$

205

Solution, cont.:



$$P(X < 1 \text{ or } Y < 1) = 1 - .375 = .625$$

Answer D

206

Exercise:

A company is reviewing tornado damage claims under a farm insurance policy. Let X be the portion of a claim representing damage to the house and let Y be the portion of the same claim representing damage to the rest of the property. The joint density function of X and Y is

$$f(x,y) = \begin{cases} 6\left[1-(x+y)\right], & x > 0, y > 0, \text{and } x+y < 1 \\ 0, & \text{otherwise} \end{cases}$$

207

Exercise, cont.:

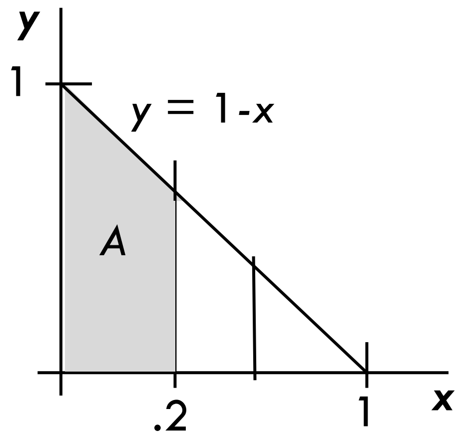
Determine the probability that the portion of a claim representing damage to the house is less than 0.2 .

(A) .360 (B) .480 (C) .488 (D) .512 (E) .520

208

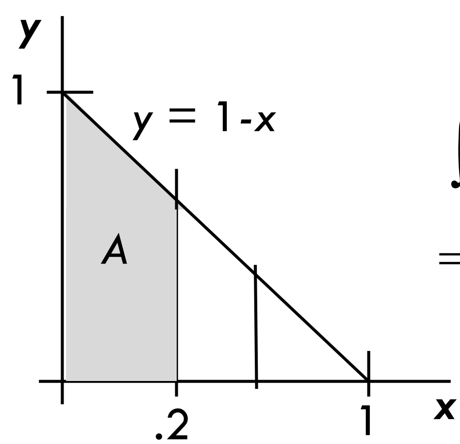
Solution:

Find $P(X<.2)=\iint_A f(x,y)\,dy\,dx$ where A is the region indicated in the diagram below.



209

Solution, cont.:



$$\iint_A f(x,y)\,dy\,dx$$

$$=6\int_0^{.2}\int_0^{1-x}\left[1-x-y\right]dy\,dx$$

210

Solution, cont.:

$$6\int_0^{.2}\int_0^{1-x}\left[1-x-y\right]dy\,dx = 6\int_0^{.2}\left(y-xy-\frac{y^2}{2}\right)\Bigg|_0^{1-x}dx$$

$$=6\int_0^{.2}\left((1-x)-x(1-x)-\frac{(1-x)^2}{2}\right)dx$$

$$=6\int_0^{.2}\frac{(1-x)^2}{2}dx = 6\left[\frac{-(1-x)^3}{6}\right]\Bigg|_0^{.2} = .488$$

Answer C

Note: We can also work this using the marginal for X. The calculations are basically the same.

211

Definitions:

- *Discrete Case*. The conditional distribution of X given that $Y=y$ is given by

$$P(X=x\mid Y=y)=p(x\mid y)=\frac{p(x,y)}{p_Y(y)}.$$

- *Continuous Case*. Let X and Y be continuous random variables with joint density function $f(x,y)$. The conditional density for X given that $Y=y$ is given by

$$f(x\mid Y=y)=f(x\mid y)=\frac{f(x,y)}{f_Y(y)}.$$

212

Conditional Expected Value:

- For discrete random variables,

$$E(Y \mid X = x) = \sum_y y\, p(y \mid x)$$

$$E(X \mid Y = y) = \sum_x x\, p(x \mid y)$$

- When X and Y are continuous, the conditional expected values are given by

$$E(Y \mid X = x) = \int_{-\infty}^{\infty} y\, f(y \mid x)\, dy$$

$$E(X \mid Y = y) = \int_{-\infty}^{\infty} x\, f(x \mid y)\, dx$$

213

Definitions:

- Two discrete random variables *X* and *Y* are independent if

$$p(x, y) \;=\; p_X(x)\, p_Y(y)$$

for all pairs of outcomes (*x*,*y*).

- Two continuous random variables *X* and *Y* are independent if

$$f(x, y) \;=\; f_X(x)\, f_Y(y)$$

for all pairs (*x*,*y*).

214

Exercise:

A diagnostic test for the presence of a disease has two possible outcomes: 1 for disease present and 0 for disease not present. Let X denote the disease state of a patient, and let Y denote the outcome of the diagnostic test.

215

Exercise, cont.:

The joint probability function of X and Y is given by:

$P(X = 0, Y = 0) = 0.800$
$P(X = 1, Y = 0) = 0.050$
$P(X = 0, Y = 1) = 0.025$
$P(X = 1, Y = 1) = 0.125$

Calculate $\text{Var}(Y \mid X = 1)$.

(A) 0.13 (B) 0.15 (C) 0.20 (D) 0.51 (E) 0.71

216

Solution:

We can calculate this variance if we know the conditional distribution of Y given that $X=1$.

$Y \backslash X$	0	1
0	.800	.050
1	.025	.125
$p_X(x)$	.825	.175

Solution, cont.:

$$P(Y=0 \mid X=1)=\frac{P(Y=0 \,\&\, X=1)}{P(X=1)}$$

$$=\frac{.05}{.175}=.2857$$

$$P(Y=1 \mid X=1)=\frac{P(Y=1 \,\&\, X=1)}{P(X=1)}$$

$$=\frac{.125}{.175}=.7143$$

Solution, cont.:

Use $V(X)=E\left(X^2\right)-E\left(X\right)^2$.

$$E\left(Y\mid X=1\right)=.2857\left(0\right)+.7143\left(1\right)=.7143$$

$$E\left(Y^2\mid X=1\right)=.2857\left(0\right)^2+.7143\left(1\right)^2=.7143$$

$$V\left(X\right)=.7143-\left(.7143\right)^2=.204$$

Answer C

219

Exercise:

Once a fire is reported to a fire insurance company, the company makes an initial estimate, *X*, of the amount it will pay to the claimant for the fire loss. When the claim is finally settled, the company pays an amount, *Y*, to the claimant. The company has determined that *X* and *Y* have the joint density function

$$f\left(x,y\right)=\frac{2}{x^2\left(x-1\right)}y^{-\frac{(2x-1)}{(x-1)}},\quad x>1,\ y>1$$

220

Exercise, cont.:

Given that the initial claim estimated by the company is 2, determine the probability that the final settlement amount is between 1 and 3.

A) 1/9 B) 2/9 C) 1/3 D) 2/3 E) 8/9

221

Solution:

To find $P\left(1 < Y < 3 \mid X = 2\right)$ we need:

$$f\left(y \mid X = 2\right) = \frac{f\left(2, y\right)}{f_X\left(2\right)} = \frac{.5y^{-3}}{f_X\left(2\right)}$$

$$f_X\left(2\right) = \int_1^{\infty} f\left(2, y\right) dy = \int_1^{\infty} .5y^{-3} dy = \left.\frac{-y^{-2}}{-4}\right|_1^{\infty} = \frac{1}{4}$$

$$f\left(y \mid X = 2\right) = \frac{.5y^{-3}}{f_X\left(2\right)} = \frac{.5y^{-3}}{\left(1/4\right)} = 2y^{-3}$$

$$P\left(1 < Y < 3 \mid X = 2\right) = \int_1^3 f\left(y \mid X = 2\right) dy$$

$$= \int_1^3 2y^{-3} dy = -y^{-2}\Big|_1^3 = \frac{8}{9}$$

Answer E

222

Review:

Counting Partitions:

The number of partitions of n objects into k distinct groups of size $n_1, n_2, \ldots, n_k$ is given by

$$\binom{n}{n_1,\ n_2, \ldots,\ n_k} = \frac{n!}{n_1!\ n_2! \ldots n_k!}$$

Review, cont.:

The Multinomial Distribution:

Random experiment has k mutually exclusive outcomes $E_1, \ldots, E_k$, with $P(E_i) = p_i$. Repeat this experiment in n independent trials.

Let X_i be the number of times that the outcome E_i occurs in the n trials.

$$P\left(X_1 = n_1 \,\&\, X_2 = n_2 \,\&\ldots\&\, X_k = n_k\right)$$

$$= \binom{n}{n_1,\ n_2, \ldots,\ n_k} p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}$$

Exercise:

A large pool of adults earning their first driver's license includes 50% low-risk drivers, 30% moderate-risk drivers, and 20% high-risk drivers. Because these drivers have no prior driving record, an insurance company considers each driver to be randomly selected from the pool. This month, the insurance company writes 4 new policies for adults earning their first driver's license.

Exercise, cont.:

What is the probability that these 4 will contain at least two more high-risk drivers than low-risk drivers?

(A) .006 (B) .012 (C) .018 (D) .049 (E) .073

Solution:

L, *M*, and *H*: number of low risk, moderate risk and high risk drivers respectively.

$$p_1 = P(L) = .50$$

$$p_2 = P(M) = .30$$

$$p_3 = P(H) = .20$$

There are four cases:

Solution, cont.:

1 $$P\left(L = 0 \,\&\, M = 0 \,\&\, H = 4\right) = \binom{4}{0,0,4} .5^0 .3^0 .2^4$$

$$= 1\left(.5^0 .3^0 .2^4\right)$$

$$= .0016$$

2 $$P\left(L = 0 \,\&\, M = 1 \,\&\, H = 3\right) = \binom{4}{0,1,3} .5^0 .3^1 .2^3$$

$$= 4\left(.5^0 .3^1 .2^3\right)$$

$$= .0096$$

Solution, cont.:

3 $P(L=1 \,\&\, M=0 \,\&\, H=3)=\binom{4}{1,0,3}.5^1.3^0.2^3$

$= 4\left(.5^1.3^0.2^3\right)$

$= .0160$

4 $P(L=0 \,\&\, M=2 \,\&\, H=2)=\binom{4}{0,2,2}.5^0.3^2.2^2$

$= 6\left(.5^0.3^2.2^2\right)$

$= .0216$

Solution, cont.:

The sum of these probabilities is .0488.

Answer D

Expected Value Properties:

- **Sum of Two Random Variables:**

$$E(X+Y) = E(X) + E(Y)$$

- **Product of Two Random Variables:**

Discrete Case:

$$E\left[(XY)\right] = \sum_x \sum_y (xy)\, p(x,y)$$

Continuous Case:

$$E\left[(XY)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f\left(x,y\right) dy\, dx$$

- **Product when X and Y Independent:**

$$E(XY) = E(X)E(Y)$$

231

Covariance Properties:

- **Covariance of X and Y:**

$$Cov(X,Y) = E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]$$

- **Alternative Calculation:**

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

- **X and Y Independent:**

$$Cov(X,Y) = 0$$

232

Variance Properties:

- **Variance of X + Y:**

$$V(X+Y) \;=\; V(X) \;+\; V(Y) \;+\; 2\mathrm{Cov}(X,Y)$$

- **Variance of X+Y when X and Y Independent:**

$$V(X+Y) \;=\; V(X) \;+\; V(Y)$$

233

Useful Properties of Covariance:

1 $\mathrm{Cov}(X,Y) \;=\; \mathrm{Cov}(Y,X)$

2 $\mathrm{Cov}(X,X) \;=\; V(X)$

3 If k is a constant random variable, then $\mathrm{Cov}(X,k) = 0$.

4 $\mathrm{Cov}(aX,bY) \;=\; ab\,\mathrm{Cov}(X,Y)$

5 $\mathrm{Cov}(X,Y+Z) \;=\; \mathrm{Cov}(X,Y) \;+\; \mathrm{Cov}(X,Z)$

234

Correlation Coefficient:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}, \quad -1 \le \rho_{X,Y} \le 1$$

Exercise:

Let X and Y be the number of hours that a randomly selected person watches movies and sporting events, respectively, during a three-month period. The following information is known about X and Y:

$$E(X) = 50 \qquad Var(X) = 50$$

$$E(Y) = 20 \qquad Var(Y) = 30$$

$$Cov(X,Y) = 10$$

Exercise, cont.:

One hundred people are randomly selected and observed for these three months. Let T be the total number of hours that these one hundred people watch movies or sporting events during this three-month period.

Approximate the value of $P(T < 7100)$.

(A) 0.62 (B) 0.84 (C) 0.87 (D) 0.92 (E) 0.97

Solution:

A) Look at the total hours for a single individual

B) Use the central limit theorem and normal approximation.

For one individual, the total hours watching movies or sporting events is $S = X + Y$.

$$E(S) = E(X+Y) = E(X) + E(Y) = 50 + 20 = 70$$

$$\begin{aligned} V(S) &= V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X,Y) \\ &= 50 + 30 + 2(10) = 100 \end{aligned}$$

Solution, cont.:

One hundred people are assumed iid. The total for all 100 people is $T = S_1 + ... + S_{100}$.

By the central limit theorem, T is approximately normal with

$$E(T) = \mu_T = 100(70) = 7,000$$

$$V(T) = \sigma_T{}^2 = 100(100) = 10,000$$

$$\sigma_S = \sqrt{10,000} = 100$$

Thus, $P\left(T < 7100\right) = P\left(Z < \dfrac{7,100 - 7,000}{100}\right)$

$$= P\left(Z < 1\right) = .8413$$

Answer B

Exercise:

Let X and Y be continuous random variables with joint density function

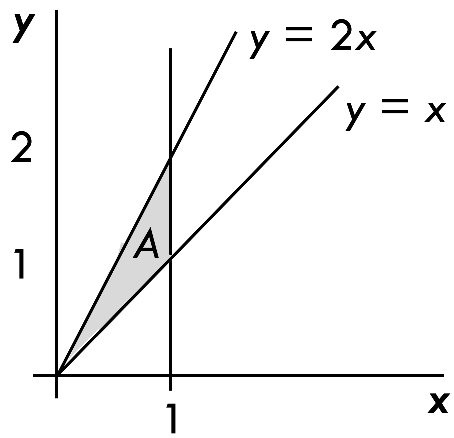
$$f\left(x,y\right) = \begin{cases} \dfrac{8}{3}xy & \text{for } 0 \le x \le 1,\ x \le y \le 2x \\ 0 & \text{otherwise} \end{cases}$$

Calculate the covariance of X and Y.

(A) 0.04 (B) 0.25 (C) 0.67 (D) 0.80 (E) 1.24

Solution:

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$$



Solution, cont.:

$$E(XY) = \iint_R xyf(x,y)dydx = \frac{8}{3}\int_0^1 \int_x^{2x} x^2y^2\,dy\,dx$$

$$= \frac{8}{3}\int_0^1 \left[x^2 \frac{y^3}{3}\Bigg|_x^{2x} \right] dx = \frac{8}{3}\int_0^1 \left(\frac{7}{3}x^5 \right) dx$$

$$= \frac{56}{9}\left(\frac{x^6}{6}\right)\Bigg|_0^1 = \frac{56}{54}$$

Solution, cont.:

$$E(Y) = \iint_R y\, f(x,y)\, dy\, dx = \frac{8}{3}\int_0^1 \int_x^{2x} xy^2 dy\, dx$$

$$= \frac{8}{3}\int_0^1 \left[x \frac{y^3}{3}\Bigg|_x^{2x} \right] dx = \frac{8}{3}\int_0^1 \left(\frac{7}{3} x^4 \right) dx$$

$$= \frac{56}{9}\left(\frac{x^5}{5} \right)\Bigg|_0^1 = \frac{56}{45}$$

243

Solution, cont.:

$$E(X) = \iint_R x\, f(x,y)\, dy\, dx = \frac{8}{3}\int_0^1 \int_x^{2x} x^2 y\, dy\, dx$$

$$= \frac{8}{3}\int_0^1 \left[x^2 \frac{y^2}{2}\Bigg|_x^{2x} \right] dx$$

$$= \frac{8}{3}\int_0^1 \left(\frac{3}{2} x^4 \right) dx$$

$$= 4\left(\frac{x^5}{5} \right)\Bigg|_0^1 = \frac{4}{5}$$

244

Solution, cont.:

$$\begin{aligned}\operatorname{Cov}(X,Y) &= E(XY)-E(X)E(Y)\\ &= \frac{56}{54}-\frac{4}{5}\left(\frac{56}{45}\right)\\ &= .041\end{aligned}$$

Answer A

Exercise:

X: Size of a surgical claim
Y: Size of the associated hospital claim

An actuary is using a model in which

$E(X) = 5$,
$E(X^2) = 27.4$,
$E(Y) = 7$,
$E(Y^2) = 51.4$, and
$V(X+Y) = 8$.

Exercise, cont.:

Let $C_1 = X+Y$ denote the size of the combined claims before the application of a 20% surcharge on the hospital portion of the claim, and let C_2 denote the size of the combined claims after the application of that surcharge. Calculate $Cov(C_1, C_2)$.

(A) 8.80 (B) 9.60 (C) 9.76 (D) 11.52 (E) 12.32

Solution:

$$C_2 = X + 1.2Y$$

$$Cov(C_1, C_2) = Cov\left(X+Y, X+1.2Y\right)$$

$$= Cov(X,X) + Cov(X,1.2Y) + Cov(Y,X) + Cov(Y,1.2Y)$$

$$= V(X) + 1.2Cov(X,Y) + Cov(X,Y) + 1.2Cov(Y,Y)$$

$$= V(X) + 2.2Cov(X,Y) + 1.2V(Y)$$

$$V\left(X\right) = E\left(X^2\right) - E(X)^2 = 27.4 - 5^2 = 2.4$$

$$V\left(Y\right) = E\left(Y^2\right) - E(Y)^2 = 51.4 - 7^2 = 2.4$$

$$V\left(X+Y\right) = V\left(X\right) + V\left(Y\right) + 2Cov\left(X,Y\right)$$

Solution, cont.:

$$8 = 2.4 + 2.4 + 2\text{Cov}(X, Y)$$

$$\text{Cov}(X, Y) = 1.6$$

Now we have the required information.

$$\begin{aligned}\text{Cov}(C_1, C_2) &= V(X) + 2.2\text{Cov}(X, Y) + 1.2V(Y)\\ &= 2.4 + 2.2(1.6) + 1.2(2.4) = 8.80\end{aligned}$$

Answer A

Theorem:

Double Expectation Theorem of the Mean:

$E[E(X \mid Y)] = E(X)$ and $E[E(Y \mid X)] = E(Y)$

Exercise:

An auto insurance company insures an automobile worth 15,000 for one year under a policy with a 1,000 deductible. During the policy year there is a 0.04 chance of partial damage to the car and a 0.02 chance of a total loss of the car.

Exercise, cont.:

If there is partial damage to the car, the amount X of damage (in thousands) follows a distribution with density function

$$f(x)=\begin{cases}0.5003e^{-x/2} & \text{for } 0<x<15\\ 0 & \text{otherwise}\end{cases}$$

What is the expected claim payment?

(A) 320 (B) 328 (C) 352 (D) 380 (E) 540

Solution:

There are three possible cases. Amounts are expressed in thousands.

a) No damage.

$$P(\text{No Damage}) = 1 - .04 - .02 = .94$$

$$E(\text{Amount Paid} \mid \text{No Damage}) = 0$$

b) Full damage.

$$P(\text{Total Loss}) = .02$$

$$E(\text{Amount Paid} \mid \text{Total Loss}) = 15 - 1 = 14$$

253

Solution, cont.:

c) Partial damage.

$$P(\text{Partial Damage}) = .04$$

$$E(\text{Amount Paid} \mid \text{Partial Damage})$$

$$= \int_1^{15} (x-1) f(x)\, dx$$

$$= .5003 \int_1^{15} (x-1) e^{-x/2} dx = 1.2049$$

Expected amount paid in thousands

$$.94(0) + .02(14) + .04(1.2049) = .328$$

Answer B

254