STAT415
probability(2)

part 2 –problems from SOA

each play. Note that this mathematical expectation can be written

$$E(X) = \sum_{x} x f(x)$$

and is often denoted by the Greek letter μ, which is called the mean of X or of its distribution.

Definition 2.2-1

If $f(x)$ is the pmf of the random variable X of the discrete type with space S, and if the summation

$$\sum_{x \in S} u(x) f(x), \quad \text{which is sometimes written} \quad \sum_{S} u(x) f(x),$$

exists, then the sum is called the **mathematical expectation** or the **expected value** of $u(X)$, and it is denoted by $E[u(X)]$. That is,

$$E[u(X)] = \sum_{x \in S} u(x) f(x).$$

Theorem 2.2-1

When it exists, the mathematical expectation E satisfies the following properties:

(a) If c is a constant, then $E(c) = c$.

(b) If c is a constant and u is a function, then

$$E[c\,u(X)] = cE[u(X)].$$

(c) If c_1 and c_2 are constants and u_1 and u_2 are functions, then

$$E[c_1u_1(X) + c_2u_2(X)] = c_1E[u_1(X)] + c_2E[u_2(X)].$$

24. The number of injury claims per month is modeled by a random variable N with

$$P[N=n]=\frac{1}{(n+1)(n+2)}$$, for nonnegative integers, n.

Calculate the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

(A) 1/3
(B) 2/5
(C) 1/2
(D) 3/5
(E) 5/6

24. Solution: B

$$P\left[N \geq 1 \middle| N \leq 4\right] = \frac{P[1 \leq N \leq 4]}{P[N \leq 4]} = \left[\frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \frac{1}{30}\right] \Big/ \left[\frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \frac{1}{30}\right]$$

$$= \frac{10+5+3+2}{30+10+5+3+2} = \frac{20}{50} = \frac{2}{5}$$

64. A probability distribution of the claim sizes for an auto insurance policy is given in the table below:

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

Calculate the percentage of claims that are within one standard deviation of the mean claim size.

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

64. Solution: A

The mean is 20(0.15) + 30(0.10) + 40(0.05) + 50(0.20) + 60(0.10) + 70(0.10) + 80(0.30) = 55. The second moment is 400(0.15) + 900(0.10) + 1600(0.05) + 2500(0.20) + 3600(0.10) + 4900(0.10) + 6400(0.30) = 3500. The variance is $3500 – 55^2 = 475$. The standard deviation is 21.79. The range within one standard deviation of the mean is 33.21 to 76.79, which includes the values 40, 50, 60, and 70. The sum of the probabilities for those values is 0.05 + 0.20 + 0.10 + 0.10 = 0.45.

44. An insurance policy pays 100 per day for up to three days of hospitalization and 50 per day for each day of hospitalization thereafter.

The number of days of hospitalization, X, is a discrete random variable with probability function

$$P[X=k]=\begin{cases}\dfrac{6-k}{15}, & k=1,2,3,4,5\\ 0, & \text{otherwise.}\end{cases}$$

Determine the expected payment for hospitalization under this policy.

(A) 123
(B) 210
(C) 220
(D) 270
(E) 367

44. Solution: C

The probabilities of 1, 2, 3, 4, and 5 days of hospitalization are 5/15, 4/15, 3/15, 2/15, and 1/15 respectively. The expected payments are 100, 200, 300, 350, and 400 respectively. The expected value is [100(5) + 200(4) + 300(3) + 350(2) + 400(1)]/15 = 220.

101. The profit for a new product is given by $Z = 3X - Y - 5$. X and Y are independent random variables with $\text{Var}(X) = 1$ and $\text{Var}(Y) = 2$.

Calculate $\text{Var}(Z)$.

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

101. Solution: D

Due to the independence of X and Y

$$Var(Z) = Var(3X - Y - 5) = 3^2 Var(X) + (-1)^2 Var(Y) = 9(1) + 2 = 11.$$

$$Var(b) = 0.$$

$$Var(aX + b) = a^2 Var(X).$$

$$Var(X \pm Y) = Var(X) + Var(Y) \pm 2cov(X, Y)$$

$$Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2$$

Definition 2.3-1

Let X be a random variable of the discrete type with pmf $f(x)$ and space S. If there is a positive number h such that

$$E(e^{tX}) = \sum_{x \in S} e^{tx} f(x)$$

exists and is finite for $-h < t < h$, then the function defined by

$$M(t) = E(e^{tX})$$

is called the **moment-generating function of** X (or of the distribution of X). This function is often abbreviated as mgf.

Setting $t = 0$, we see that

$$M'(0) = \sum_{x \in S} x f(x) = E(X),$$

$$M''(0) = \sum_{x \in S} x^2 f(x) = E(X^2),$$

and, in general,

$$M^{(r)}(0) = \sum_{x \in S} x^r f(x) = E(X^r).$$

In particular, if the moment-generating function exists, then

$$M'(0) = E(X) = \mu \qquad \text{and} \qquad M''(0) - [M'(0)]^2 = E(X^2) - [E(X)]^2 = \sigma^2.$$

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

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

Calculate 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

57. Solution: B

This is the moment generating function of a gamma distribution with parameters 4 and 2,500. The standard deviation is the square root of the shape parameter times the scale parameter, or 2(2,500) = 5,000. But such recognition is not necessary.

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

$$M''(t) = 20(2500)^2(1-2500t)^{-6}$$

$$E(X) = M'(0) = 10,000$$

$$E(X^2) = M''(0) = 125,000,000$$

$$Var(X) = 125,000,000 - 10,000^2 = 25,000,000$$

$$SD(X) = 5,000$$

58. 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 mutually independent.

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

$$M_J(t) = (1-2t)^{-3}, \quad M_K(t) = (1-2t)^{-2.5}, \quad M_L(t) = (1-2t)^{-4.5}.$$

Let X represent the combined losses from the three cities.

Calculate $E(X^3)$.

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

58. Solution: E

Because the losses are independent, the mgf of their sum is the product of the individual mgfs, which is $(1 - 2t)^{-10}$. The third moment can be determined by evaluating the third derivative at zero. This is $(10)(2)(11)(2)(12)(2)(1- 2(0))^{-13} = 10{,}560$.

Joint moment generating functions

Definition

Let $(X_1, X_2, \ldots, X_n)$ be a random vector. The *joint moment generating function* of $(X_1, X_2, \ldots, X_n)$ is

$$M(t_1, t_2, \ldots, t_n) = E\left[e^{t_1X_1 + t_2X_2 + \cdots + t_nX_n}\right]$$

95. X and Y are independent random variables with common moment generating function $M(t) = \exp(t^2/2)$.

Let $W = X + Y$ and $Z = Y - X$.

Determine the joint moment generating function, $M(t_1, t_2)$ of W and Z.

(A) $\exp(2t_1^2 + 2t_2^2)$

(B) $\exp[(t_1 - t_2)^2]$

(C) $\exp[(t_1 + t_2)^2]$

(D) $\exp(2t_1 t_2)$

(E) $\exp(t_1^2 + t_2^2)$

95. Solution: E

$$M(t_1,t_2) = E\left[e^{t_1W+t_2Z}\right] = E\left[e^{t_1(X+Y)+t_2(Y-X)}\right] = E\left[e^{(t_1-t_2)X}e^{(t_1+t_2)Y}\right]$$

$$= E\left[e^{(t_1-t_2)X}\right]E\left[e^{(t_1+t_2)Y}\right] = e^{\frac{1}{2}(t_1-t_2)^2}e^{\frac{1}{2}(t_1+t_2)^2} = e^{\frac{1}{2}\left(t_1{}^2-2t_1t_2+t_2^2\right)}e^{\frac{1}{2}\left(t_1^2+2t_1t_2+t_2^2\right)} = e^{t_1{}^2+t_2{}^2}.$$

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

$$p(x) = \begin{cases} 1/3, & x = 0 \\ 2/3, & x = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Determine the moment generating function, $M(t)$, of $Y = X_1X_2X_3$.

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

(B) $1 + 2e^t$

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

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

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

98. Solution: A

The product of the three variables is 1 only if all three are 1, so $P[Y=1]=8/27$. The remaining probability of 19/27 is on the value 0. The mgf is

$$M(t)=E[e^{tY}]=e^{t(0)}(19/27)+e^{t(1)}(8/27)=\frac{19}{27}+\frac{8}{27}e^{t}.$$

130. The value of a piece of factory equipment after three years of use is $100(0.5)^X$ where X is a random variable having moment generating function

$$M_X(t) = \frac{1}{1-2t}, \quad t < \frac{1}{2}.$$

Calculate the expected value of this piece of equipment after three years of use.

(A) 12.5
(B) 25.0
(C) 41.9
(D) 70.7
(E) 83.8

130. Solution: C

$$E[100(0.5)^X] = 100E[e^{X\ln 0.5}] = 100M_X(\ln 0.5) = 100\frac{1}{1-2\ln 0.5} = 41.9.$$

165. Two claimants place calls simultaneously to an insurer's claims call center. The times X and Y, in minutes, that elapse before the respective claimants get to speak with call center representatives are independently and identically distributed. The moment generating function of each random variable is

$$M(t) = \left(\frac{1}{1-1.5t}\right)^2, \quad t < \frac{2}{3}.$$

Calculate the standard deviation of $X + Y$.

(A) 2.1
(B) 3.0
(C) 4.5
(D) 6.7
(E) 9.0

165. Solution: B

For either distribution the moments can be found from

$$M(t) = (1-1.5t)^{-2}$$

$$M'(t) = 2(1.5)(1-1.5t)^{-3} = 3(1-1.5t)^{-3}$$

$$M''(t) = 3(3)(1.5)(1-1.5t)^{-4} = 13.5(1-1.5t)^{-4}$$

$$E(X) = E(Y) = M'(0) = 3$$

$$E(X^2) = E(Y^2) = M''(0) = 13.5$$

$$Var(X) = Var(Y) = 13.5 - 3^2 = 4.5$$

$$Var(X+Y) = Var(X) + Var(Y) = 4.5 + 4.5 = 9.$$

The standard deviation is the square root, 3.

If it is recognized that this is the moment generating function of the gamma distribution, then the parameters (1.5 and 2) and the moments can be obtained without calculations as $1.5(2) = 3$ and $1.5(1.5)(3) = 4.5$.

222. Let X represent the number of policies sold by an agent in a day. The moment generating function of X is

$$M(t) = 0.45e^{t} + 0.35e^{2t} + 0.15e^{3t} + 0.05e^{4t}, \quad \text{for } -\infty < t < \infty.$$

Calculate the standard deviation of X.

(A) 0.76
(B) 0.87
(C) 1.48
(D) 1.80
(E) 4.00

222. Solution: B

One approach is to take derivatives of the mgf and set them equal to zero. This yields a mean of 0.45 + 0.35(2) + 0.15(3) + 0.05(4) = 1.8 and a second moment of 0.45 + 0.35(4) + 0.15(9) + 0.05(16) = 4. The variance is 4 – 3.24 = 0.76 and the standard deviation is 0.87.

Alternatively, it can be recognized that this mgf corresponds to a discrete random variable with probabilities 0.45, 0.35, 0.15, and 0.05 at 1, 2, 3, and 4, respectively. The same formulas result.

Suppose that an urn contains N_1 success balls and N_2 failure balls. Let $p = N_1/(N_1 + N_2)$, and let X equal the number of success balls in a random sample of size n that is taken from this urn. If the sampling is done one at a time with replacement, then the distribution of X is $b(n, p)$; if the sampling is done without replacement, then X has a hypergeometric distribution with pmf

$$f(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N_1+N_2}{n}},$$

where x is a nonnegative integer such that $x \le n$, $x \le N_1$, and $n - x \le N_2$. When $N_1 + N_2$ is large and n is relatively small, it makes little difference if the sampling is done with or without replacement. In Figure 2.4-4, the probability histograms are compared for different combinations of n, N_1, and N_2.

132. A store has 80 modems in its inventory, 30 coming from Source A and the remainder from Source B. Of the modems from Source A, 20% are defective. Of the modems from Source B, 8% are defective.

Calculate the probability that exactly two out of a sample of five modems selected without replacement from the store's inventory are defective.

(A) 0.010
(B) 0.078
(C) 0.102
(D) 0.105
(E) 0.125

132. Solution: C

The number of defective modems is 20% x 30 + 8% x 50 = 10.

The probability that exactly two of a random sample of five are defective is $\dfrac{\binom{10}{2}\binom{70}{3}}{\binom{80}{5}} = 0.102$.

151. From 27 pieces of luggage, an airline luggage handler damages a random sample of four.

The probability that exactly one of the damaged pieces of luggage is insured is twice the probability that none of the damaged pieces are insured.

Calculate the probability that exactly two of the four damaged pieces are insured.

(A) 0.06
(B) 0.13
(C) 0.27
(D) 0.30
(E) 0.31

151. Solution: C

The ratio of the probability that one of the damaged pieces is insured to the probability that none of the damaged pieces are insured is

$$\frac{\dfrac{\binom{r}{1}\binom{27-r}{3}}{\binom{27}{4}}}{\dfrac{\binom{r}{0}\binom{27-r}{4}}{\binom{27}{4}}} = \frac{4r}{24-r},$$

where r is the total number of pieces insured. Setting this ratio equal to 2 and solving yields $r = 8$.

The probability that two of the damaged pieces are insured is

$$\frac{\binom{r}{2}\binom{27-r}{2}}{\binom{27}{4}} = \frac{\binom{8}{2}\binom{19}{2}}{\binom{27}{4}} = \frac{(8)(7)(19)(18)(4)(3)(2)(1)}{(27)(26)(25)(24)(2)(1)(2)(1)} = \frac{266}{975} = 0.27.$$

170. An insurance agent meets twelve potential customers independently, each of whom is equally likely to purchase an insurance product. Six are interested only in auto insurance, four are interested only in homeowners insurance, and two are interested only in life insurance.

The agent makes six sales.

Calculate the probability that two are for auto insurance, two are for homeowners insurance, and two are for life insurance.

(A) 0.001
(B) 0.024
(C) 0.069
(D) 0.097
(E) 0.500

170. Solution: D

$$\frac{\binom{6}{2}\binom{4}{2}\binom{2}{2}}{\binom{12}{6}} = \frac{15(6)(1)}{924} = 0.097$$

177. In a group of 25 factory workers, 20 are low-risk and five are high-risk.

Two of the 25 factory workers are randomly selected without replacement.

Calculate the probability that exactly one of the two selected factory workers is low-risk.

(A) 0.160
(B) 0.167
(C) 0.320
(D) 0.333
(E) 0.633

177. Solution: D

This is a hypergeometric probability,

$$\frac{\binom{20}{1}\binom{5}{1}}{\binom{25}{2}} = \frac{20(5)}{25(24)/2} = \frac{100}{300} = 0.333 \ ,$$

Alternatively, the probability of the first worker being high risk and the second low risk is (5/25)(20/24) = 100/600 and of the first being low risk and the second high risk is (20/25)(5/24) = 100/600 for a total probability of 200/600 = 0.333.

210. On a block of ten houses, k are not insured. A tornado randomly damages three houses on the block.

The probability that none of the damaged houses are insured is 1/120.

Calculate the probability that at most one of the damaged houses is insured.

(A) 1/5
(B) 7/40
(C) 11/60
(D) 49/60
(E) 119/120

210. Solution: C

The probability that none of the damaged houses are insured is

$$\frac{1}{120} = \frac{\binom{10-k}{0}\binom{k}{3}}{\binom{10}{3}} = \frac{k(k-1)(k-2)}{720}.$$

$k(k-1)(k-2) = 6$

This cubic equation could be solved by expanding, subtracting 6, and refactoring. However, because k must be an integer, the three factors must be integers and thus must be 3(2)(1) for $k =$ 3.

The probability that at most one of the damaged houses is insured equals

$$\frac{1}{120} + \frac{\binom{10-3}{1}\binom{3}{2}}{\binom{10}{3}} = \frac{1}{120} + \frac{7(3)}{120} = \frac{22}{120} = \frac{11}{60}.$$

211. In a casino game, a gambler selects four different numbers from the first twelve positive integers. The casino then randomly draws nine numbers without replacement from the first twelve positive integers. The gambler wins the jackpot if the casino draws all four of the gambler's selected numbers.

Calculate the probability that the gambler wins the jackpot.

(A) 0.002
(B) 0.255
(C) 0.296
(D) 0.573
(E) 0.625

211. Solution: B

This question is equivalent to "What is the probability that 9 different chips randomly drawn from a box containing 4 red chips and 8 blues chips will contain the 4 red chips?" The hypergeometric probability is

$$\frac{\binom{4}{4}\binom{8}{5}}{\binom{12}{9}} = \frac{1(56)}{220} = 0.2545.$$