2.5 Mean and Variance

The *mean* (or *expected value*) of a discrete random variable X with range R and pmf $f(x)$ is, provided series converges absolutely,

$$\mu = E(X) = \sum_{x \in R} x f(x).$$

The *variance* of discrete random variable X with range R and pmf $f(x)$ is, provided series converges absolutely,

$$\begin{aligned} \sigma^2 = Var(X) &= \sum_{x \in R} (x-\mu)^2 f(x) = E[(X-\mu)^2] \\ &= \sum_{x \in R} x^2 f(x) - \mu^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2, \end{aligned}$$

with associated *standard deviation*, σ.

For *uniform* random variable X with range $R = \{1, 2, \ldots, k\}$, and pmf $f(x) = \frac{1}{k}$,

$$\mu = E(X) = \frac{k+1}{2}, \quad \sigma^2 = Var(X) = \frac{k^2-1}{12}, \quad \sigma = \sqrt{\frac{k^2-1}{12}};$$

for *binomial* random variable,

$$\mu = E(X) = np, \quad \sigma^2 = Var(X) = npq, \quad \sigma = \sqrt{npq};$$

for *Poisson* random variable,

$$\mu = E(X) = \lambda, \quad \sigma^2 = Var(X) = \lambda, \quad \sigma = \sqrt{\lambda}.$$

Exercise 2.5 (Mean and Variance)

1. *Expected value, variance and standard deviation: seizures.* The probability function for the number of seizures, X, of a typical epileptic person in any given year is given in the following table.

x	0	2	4	6	8	10
$f(x)$	0.17	0.21	0.18	0.11	0.16	0.17

   (a) *Calculating the expected value.* The expected value (mean) number of seizures is given by

$$E(X) = \sum_x x f(x) = 0(0.17)+2(0.21)+4(0.18)+6(0.11)+8(0.16)+10(0.17) =$$

   (i) **4.32** (ii) **4.78** (iii) **5.50** (iv) **5.75**.

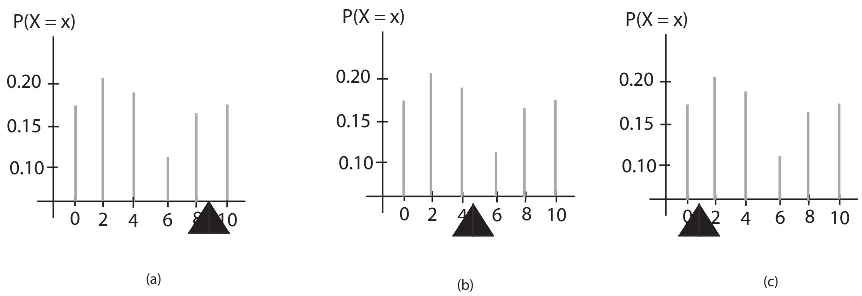

Figure 2.11: Expected value: fulcrum point of balance

```
x <- c(0,2,4,6,8,10) # values of random variable
px <- c(0.17,0.21,0.18,0.11,0.16,0.17) # probabilities
EX <- sum(x*px); EX # expected value

[1] 4.78
```

(b) *Understanding expected value: seizures.*
If the expected value is like a fulcrum point which *balances* the "weight" of the probability distribution, then the expected value is most likely close to the point of the fulcrum given in which of the three graphs above?
(i) **(a) 9** (ii) **(b) 5** (iii) **(c) 1**.

(c) *Variance.* The variance in number of seizures is given by

$$\begin{aligned}\sigma^2 = Var[X] &= E\left[(X-\mu)^2\right] \\ &= \sum_x (X-\mu)^2 f(x) \\ &= (0-4.78)^2(0.17) + (2-4.78)^2(0.21) + \cdots + (10-4.78)^2(0.17) \approx\end{aligned}$$

(i) **7.32** (ii) **8.78** (iii) **10.50** (iv) **12.07**.

```
VarX <- sum((x-EX)^2*px); VarX # variance

[1] 12.0716
```

(d) *Standard Deviation.* The standard deviation in the number of seizures is

$$\sigma = \sqrt{Var(X)} \approx \sqrt{12.07} \approx$$

(circle one) (i) **3.47** (ii) **4.11** (iii) **5.07** (iv) **6.25**.

```
SDX <- sqrt(VarX); SDX # standard deviation

[1] 3.474421
```

In other words, we expect to see about $\mu \pm \sigma = 4.78 \pm 3.47$ seizures according to the probability distribution given here.

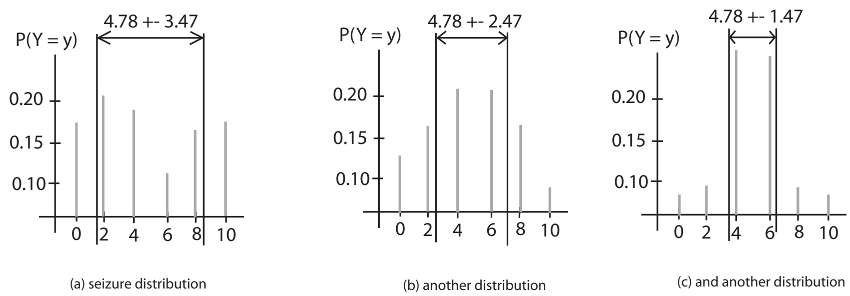

Figure 2.12: Standard deviation: dispersion of distribution

(e) *Understanding standard deviation: "dispersion".*
The standard deviation measures the dispersion of a probability distribution. The most dispersed distribution occurs in
(i) **(a)** (ii) **(b)** (iii) **(c)**.

2. *Variance and standard deviation: rolling a pair of dice.* If the dice are fair, the distribution of X (the sum of two rolls of a pair of dice) is

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

(a) *Expected value,* $g(x) = x$,

$$\mu = E(X) = \sum_x x f(x) = 2\left(\frac{1}{36}\right) + \cdots + 12\left(\frac{1}{36}\right) =$$

(i) **4** (ii) **5** (iii) **6** (iv) **7**.

```
x <- 2:12 # values of random variable
px <- c(1,2,3,4,5,6,5,4,3,2,1)/36 # probabilities
EX <- sum(x*px); EX # E(X)
```

```
[1] 7
```

(b) If $g(x) = x^2$,

$$E(X^2) = \sum_x x^2 f(x) = 2^2\left(\frac{1}{36}\right) + \cdots + 12^2\left(\frac{1}{36}\right) =$$

(i) **35.43** (ii) **47.61** (iii) **54.83** (iv) **65.67**.

```
EX2 <- sum(x^2*px); EX2 # E(X^2)
```

```
[1] 54.83333
```

(c) *Variance.*

$$\sigma^2 = V[X] = E\left[(X-\mu)^2\right] = E\left(X^2\right) - \mu^2 = 54.83 - 7^2 \approx$$

(i) **3.32** (ii) **5.83** (iii) **7.50** (iv) **8.07**.

```
VarX <- EX2 - EX^2; VarX # variance
```

```
[1] 5.833333
```

(d) *Standard deviation.*

$$\sigma = \sqrt{Var(X)} \approx \sqrt{5.83} \approx$$

(i) **2.42** (ii) **3.11** (iii) **4.07** (iv) **5.15**.

```
SDX <- sqrt(VarX); SDX # standard deviation
```

```
[1] 2.415229
```

3. *Uniform: die.* Fair die has the following uniform pmf.

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

(a) Since pmf of uniform is $f(x) = \frac{1}{k}$, in this case,

$$k =$$

(i) **5** (ii) **6** (iii) **7** (iv) **8**.

(b) The *expected* value of die is

$$\mu = E(X) = \sum_x xf(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} =$$

(i) **2.5** (ii) **3.0** (iii) **3.5** (iv) **4.0**,

```
x <- 1:6 # values of random variable
px <- rep(1/6,6) # probabilities: 1/6 repeated 6 times
EX <- sum(x*px); EX # E(X)
```

```
[1] 3.5
```

or, using the formula, the expected value of die is

$$\mu = \frac{k+1}{2} = \frac{6+1}{2} =$$

(i) **2.5** (ii) **3.0** (iii) **3.5** (iv) **4.0**.

(c) If $\mu = 11$, then

$$\mu = 11 = \frac{k+1}{2}$$

so $k =$ (i) **11** (ii) **21** (iii) **22** (iv) **23**
that is, the die has 21 sides.

4. *Another die question.* Fair six-sided die is labelled in one of three ways: there are two sides labelled 1, three sides labelled 2 and one side labelled 3. If it costs \$1 to play and you win \$1 $\times$ result from die, what is the expected value of this game?

die	1	2	3
x, payoff	$1-1$	$2-1$	$3-1$
$f(x)$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{1}{6}$

The expected value is

$$\mu = E(X) = \sum_x x f(x) = 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} + 2 \cdot \frac{1}{6} =$$

(i) $\mathbf{\frac{1}{6}}$ (ii) $\mathbf{\frac{3}{6}}$ (iii) $\mathbf{\frac{5}{6}}$ (iv) $\mathbf{\frac{7}{6}}$.

```
x <- 0:2 # values of random variable
px <- c(2,3,1)/6 # probabilities
EX <- sum(x*px); EX # E(X)
```

```
[1] 5/6
```

5. *Binomial: Airplane engines.* Each engine of four ($n = 4$) on an airplane fails 11% ($p = 0.11, q = 1 - p = 0.89$) of the time. Assume this problem obeys the conditions of a binomial experiment, in other words, X is $b(4, 0.11)$.

x	0	1	2	3	4
$f(x)$		0.310	0.058	0.005	0.000

(a) Fill in the blank: the chance no (zero) engines fail is
$f(0) = \binom{4}{0} 0.11^0 0.89^4 =$ (i) **0.005** (ii) **0.058** (iii) **0.310** (iv) **0.627**.

```
dbinom(0,4,0.11) # binomial pmf
```

```
[1] 0.6274224
```

(b) The *expected* number of failures is

$$\mu = E(X) = \sum_x x f(x) = 0(0.627)+1(0.310)+2(0.058)+3(0.005)+4(0.000) \approx$$

(i) **0.44** (ii) **0.51** (iii) **0.62** (iv) **0.73**.

or, using the formula, the expected number of failures is

$$\mu = np = 4(0.11) =$$

(i) **0.44** (ii) **0.51** (iii) **0.62** (iv) **0.73**.

(c) The *variance* in number of failures is

$$\begin{aligned}\sigma^2 = V[X] &= E\left[(X-\mu)^2\right] \\ &= \sum_x (X-\mu)^2 f(x) \\ &= (0-0.44)^2(0.627) + (1-0.44)^2(0.0.310) + \cdots + (4-0.44)^2(0.000) \approx\end{aligned}$$

(i) **0.15** (ii) **0.39** (iii) **0.51** (iv) **0.63**.

or, using the formula, the variance in number of failures is

$$\sigma^2 = npq = 4(0.11)(1-0.11) \approx$$

(i) **0.15** (ii) **0.39** (iii) **0.51** (iv) **0.63**.

(d) The *standard deviation* in number of failures is
$\sigma = \sqrt{0.39} \approx$ (i) **0.45** (ii) **0.56** (iii) **0.63** (iv) **0.83**.

6. *Bernoulli: mean and variance formulas.* Bernoulli pmf is given by:

x	0	1
$f(x)$	$1-p$	p

(a) The *expected value* is

$$\mu = E(X) = \sum_x x f(x) = 0(1-p) + 1(p) =$$

(i) $\boldsymbol{p}$ (ii) $\boldsymbol{1-p}$ (iii) $\boldsymbol{(1-p)}$ (iv) $\boldsymbol{p(1-p)}$.

(b) The *variance* is

$$\sigma^2 = Var(X) = \sum_x (x-\mu)^2 f(x) = (0-p)^2(1-p) + (1-p)^2(p) =$$

(i) $\boldsymbol{p}$ (ii) $\boldsymbol{1-p}$ (iii) $\boldsymbol{(1-p)}$ (iv) $\boldsymbol{p(1-p)}$.

7. *Poisson: accidents.* An average of $\lambda = 3$ accidents per year occurs along the I–95 stretch of highway between Michigan City, Indiana, and St. Joseph, Michigan.

(a) *Expectation.* The expected number of accidents is
$\mu = E(X) = \lambda =$
(i) **1** (ii) **2** (iii) **3** (iv) **4**.

(b) *Expected cost.* If it costs \$500,000 per accident, the expected yearly cost is
$E(C) = E(500000X) = 500000E(X) = 500000(3) =$
(i) **\$500,000** (ii) **\$1,000,000** (iii) **\$1,500,000** (iv) **\$2,000,000**.

(c) *Variance.* The variance in the number of accidents per year is
$\sigma^2 = Var(X) = \lambda =$
(i) **1** (ii) **2** (iii) **3** (iv) **4**.

(d) *Standard deviation.* Standard deviation in number of accidents per year
$\sigma = \sqrt{\lambda} = \sqrt{3} \approx$
(i) **1.01** (ii) **1.34** (iii) **1.73** (iv) **1.96**.

2.6 Functions of a Random Variable

Expected value of a function u of random variable X, $E[u(X)]$, is

$$E[u(X)] = \sum_x u(x) f(X).$$

Some properties are

$$\begin{aligned} E(a) &= a \sum_x f(x) \; = \; a, \\ E[au(X)] &= aE[u(X)], \\ E[a_1 u_1(X) + a_2 u_2(X) + \cdots + a_k u_k(X)] &= a_1 E[u_1(X)] + a_2 E[u_2(X)] + \cdots + a_k E[u_k(X)] \end{aligned}$$

where $a, a_1, a_2, \ldots a_k$ are constants (numbers, not random variables). Furthermore,

$$Var[aX] = a^2 Var(X), \quad Var[a] = 0.$$

Exercise 2.6 (Functions of a Random Variable)

1. *Functions of random value: seizures.* The pmf for the number of seizures, X, of a typical epileptic person in any given year is given in the following table.

x	0	2	4	6	8	10
$f(x)$	0.17	0.21	0.18	0.11	0.16	0.17

(a) Recall, the expected value number of seizures is

$$E(X) = \sum_x x f(x) = 0(0.17)+2(0.21)+4(0.18)+6(0.11)+8(0.16)+10(0.17) =$$

(i) **4.32** (ii) **4.78** (iii) **5.50** (iv) **5.75**.

```
x <- c(0,2,4,6,8,10) # values of random variable
px <- c(0.17,0.21,0.18,0.11,0.16,0.17) # probabilities
EX <- sum(x*px); EX # E(X)

[1] 4.78
```

(b) If the medical costs for each seizure, X, is \$200; in other words, function $u(x) = 200x$, the probability distribution for $u(X)$ is:

x	0	2	4	6	8	10
$u(x) = 200x$	$200(0) = 0$	$200(2) = 400$	800	1200	1600	2000
$p(u(x))$	0.17	0.21	0.18	0.11	0.16	0.17

The expected value (mean) cost of seizures is then given by

$$E[u(X)] = E[200X] = \sum_x (200x)f(x) = [0](0.17)+[400](0.21)+\cdots+[2000](0.17) =$$

(i) **432** (ii) **578** (iii) **750** (iv) **956**.

```
u <- 200*x # U(X), cost of seizures
EU <- sum(u*px); EU # E(U)
```

```
[1] 956
```

Alternatively, since $c = 200$ is a constant,

$$E[u(X)] = E[200X] = 200E[X] = 200(4.78) =$$

(i) **432** (ii) **578** (iii) **750** (iv) **956**.

```
EU <- 200*EX; EU # E(U)
```

```
[1] 956
```

(c) If the medical costs for each seizure, X, is given by function $u(x) = x^2$,

x	0	2	4	6	8	10
$u(x) = x^2$	$0^2 = 0$	4	16	36	64	100
$p(u(x))$	0.17	0.21	0.18	0.11	0.16	0.17

The expected value (mean) cost of seizures in this case is given by

$$E[u(X)] = E[X^2] = \sum_x x^2 f(x) = [0](0.17) + [4](0.21) + \cdots + [100](0.17) =$$

(i) **34.92** (ii) **57.83** (iii) **75.01** (iv) **94.56**.

```
EX2 <- sum(x^2*px); EX2 # E(X^2)
```

```
[1] 34.92
```

(d) If $u(x) = 200x^2 + x - 5$,

$$\begin{aligned} E[u(X)] &= E[200X^2 + X - 5] \\ &= E(200X^2) + E(X) - E(5) \\ &= 200E(X^2) + E(X) - E(5) \\ &= 200(34.92) + 4.78 - 5 \ = \end{aligned}$$

(i) **4320.67** (ii) **5780.11** (iii) **6983.78** (iv) **8480.99**.

```
u <- 200*x^2 + x - 5 # U(X), cost of seizures
EU <- sum(u*px); EU # E(U)
```

```
[1] 6983.78
```

2. *More functions of random variable: flipping until a head comes up.* A (weighted) coin has a probability of $p = 0.7$ of coming up heads (and so a probability of $1 - p = 0.3$ of coming up tails). This coin is flipped until a head comes up or until a total of 4 flips are made. Let X be the number of flips. Recall,

x	1	2	3	4
$f(x)$	0.700	0.210	0.063	0.027

   (a) If $u(X) = x$,

$$\mu = E(X) = \sum_x xf(x) = 1(0.700) + 2(0.210) = 3(0.063) + 4(0.027) =$$

   (i) **1.117** (ii) **1.217** (iii) **1.317** (iv) **1.417**.

```
x <- 1:4 # values of random variable
px <- c(0.700,0.210,0.063,0.027) # probabilities
EX <- sum(x*px); EX # E(X)
```

```
[1] 1.417
```

   (b) If $u(X) = \frac{1}{x}$,

$$E\left(\frac{1}{X}\right) = \sum_x \frac{1}{x} f(x) = \frac{1}{1}(0.7) + \frac{1}{2}(0.21) + \frac{1}{3}(0.063) + \frac{1}{4}(0.027) =$$

   (i) **0.41755** (ii) **0.83275** (iii) **1.53955** (iv) **2.56775**.

```
u <- 1/x # U(X)
EU1 <- sum(u*px); EU1 # E(U)
```

```
[1] 0.83275
```

   (c) If $u(X) = \frac{200}{x} + \frac{1}{200x} + 5$,

$$\begin{aligned} E[u(X)] &= E\left[\frac{200}{X} + \frac{1}{200X} + 5\right] \\ &= E\left[\left(200 + \frac{1}{200}\right)\frac{1}{X} + 5\right] \\ &= \left(200 + \frac{1}{200}\right) E\left[\frac{1}{X}\right] + E[5] \\ &= \left(200 + \frac{1}{200}\right)(0.83275) + 5 \ = \end{aligned}$$

   (i) **43.20** (ii) **57.80** (iii) **109.35** (iv) **171.55**.

```
u <- 200/x + 1/(200*x) + 5 # U(X)
EU <- sum(u*px); EU # E(U)
EU <- (200 + 1/200)*EU1 + 5; EU # E(U) again
```

```
[1] 171.5542
```

3. *Another function.* Assume random variable X has uniform pmf

$$f(x) = \frac{1}{4}, \quad x = 1, 2, 3, 4.$$

(a) If $u(x) = x$,

$$\mu = E(X) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{10}{4} =$$

(i) **1.0** (ii) **1.5** (iii) **2.0** (iv) **2.5**

```
x <- 1:4 # values of random variable
px <- rep(1/4,4) # probabilities: 1/4 repeated 4 times
EX <- sum(x*px); EX # E(X)
```

```
[1] 2.5
```

(b) If $u(x) = x^2$,

$$E[X^2] = 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{4} + 3^2 \cdot \frac{1}{4} + 4^2 \cdot \frac{1}{4} = \frac{30}{4} =$$

(i) **6.75** (ii) **7.00** (iii) **7.25** (iv) **7.50**

```
EX2 <- sum(x^2*px); EX2 # E(X^2)
```

```
[1] 7.5
```

and also

$$Var(X) = E(X^2) - \mu^2 = 7.5 - 2.5^2 =$$

(i) **1.25** (ii) **1.50** (iii) **1.75** (iv) **2.50**

```
VarX <- EX2 - EX^2; VarX # variance
```

```
[1] 1.25
```

(c) If $u(x) = 2x$,

$$E[2X] = 2E(X) = 2(2.5) =$$

(i) **5** (ii) **6** (iii) **7** (iv) **8**

```
u <- 2*x # U(X)
EU <- sum(u*px); EU # E(U)
```

```
[1] 5
```

and also

$$Var[2X] = 2^2 Var(X) = 4(1.25) =$$

(i) **5** (ii) **6** (iii) **7** (iv) **8**

```
VarU <- 2^2*VarX; VarU # Var(U)
```

```
[1] 5
```

(d) If $u(x) = 2x + 3x^2$,

$$E[2X + 3X^2] = 2E(X) + 3E(X^2) = 2(2.5) + 3(7.5) =$$

(i) **23.5** (ii) **24.5** (iii) **25.5** (iv) **27.5**.

```
u <- 2*x + 3*x^2 # another U(X)
EU <- 2*EX + 3*EX2; EU # E(U)
```

```
[1] 27.5
```

4. Random variable X has mean $\mu_X = \mu$, and variance $\sigma_X^2 = \sigma^2$. If $u(x) = 3x + 4$, then

$$\mu_U = E[3X + 4] = 3E(X) + 4 =$$

(i) $\boldsymbol{\mu + 4}$ (ii) $\boldsymbol{2\mu + 4}$ (iii) $\boldsymbol{3\mu + 4}$ (iv) $\boldsymbol{4\mu + 4}$

and also

$$\sigma_U^2 = Var[3X + 4] = 3^2 Var(X) + 0 =$$

(i) $\boldsymbol{8\sigma^2 + 4}$ (ii) $\boldsymbol{8\sigma^2}$ (iii) $\boldsymbol{9\sigma^2}$ (iv) $\boldsymbol{10\sigma^2}$

5. Consider random variable X where $E[X + 2] = 4$ and $E[X^2 + 4X] = 3$, then

$$E[X + 2] = E(x) + 2 = 4,$$

so $\mu = E(X) =$ (i) **1** (ii) **2** (iii) **3** (iv) **4**

and also since

$$E[X^2 + 4X] = E(X^2) + 4E(X) = E(X^2) + 4(2) = 3$$

then $E(X^2) =$ (i) $\mathbf{-5}$ (ii) $\mathbf{-3}$ (iii) $\mathbf{-1}$ (iv) **1**

so $Var(X) = E(X^2) - \mu^2 =$ (i) $\mathbf{-9}$ (ii) $\mathbf{-7}$ (iii) $\mathbf{-6}$ (iv) $\mathbf{-5}$

6. For Poisson random variable X where $\lambda = 3$, determine $E[X^2 - 3X]$.

Since

$$\mu = E[X] = \lambda =$$

(i) **1** (ii) **2** (iii) **3** (iv) **4**

and also since $Var(X) = E(X^2) - \mu^2 = \lambda$, then

$$E(X^2) = \lambda + \mu^2 = \lambda + \lambda^2 = 3 + 3^2 =$$

(i) **9** (ii) **10** (iii) **12** (iv) **13**

so

$$E[X^2 - 3X] = E(X^2) - 3E(X) = 12 - 3(3) =$$

(i) **3** (ii) **4** (iii) **5** (iv) **6**

2.7 The Moment-Generating Function

The *moment generating function (mgf)* of random variable X (taken about the origin) with pmf $f(x)$ and range R is defined by, assuming the expectation exists,

$$M(t) = E\left(e^{tX}\right) = \sum_{x \in R} e^{tX} f(x).$$

Furthermore, if random variable X and its mgf $M(t)$ exists for all t in an open interval containing 0, then

- $M(t)$ uniquely determines the distribution of X,
- $M'(0) = E(X)$, $M''(0) = E(X^2)$.

Also, if $Y = aX + b$,

$$M_Y(t) = E[e^{Yt}] = E\left[e^{(aX+b)t}\right] = E\left[e^{(at)X}e^{bt}\right] = e^{bt}E\left[e^{(at)X}\right],$$

but $E[e^{(at)X}] = M_X(at)$, so

$$M_Y(t) = e^{bt}M_X(at).$$

The *probability-generating function (pgf)* is

$$P(t) = E[t^X] = \sum_{x=0}^{\infty} t^x f(x),$$

where $P'(1) = E(X)$.

DISCRETE	$f(x)$	$M(t)$	μ	σ^2
Binomial	$\binom{n}{x} p^x q^{n-x}$	$(pe^t + q)^n$	np	npq
Poisson	$e^{-\lambda}\frac{\lambda^x}{x!}$	$e^{\lambda(e^t-1)}$	λ	λ
Geometric	$q^{x-1}p$	$\frac{pe^t}{1-qe^t}$	$1/p$	q/p^2
Negative Binomial	$\binom{x-1}{r-1} p^r q^{x-r}$	$\left(\frac{pe^t}{1-qe^t}\right)^r$	r/p	rq/p^2

Exercise 2.7 (The Moment-Generating Function)

1. *Deriving mgf from pmf.* If random variable X has range $R = \{1, 2, 3\}$ with pmf $f(1) = \frac{1}{2}$, $f(2) = \frac{1}{3}$ and $f(3) = \frac{1}{6}$, then

$$M(t) = E\left(e^{tX}\right) = \sum_{x \in R} e^{tX} f(x) =$$

(i) $\boldsymbol{\frac{1}{3}e^t + \frac{1}{2}e^{2t} + \frac{1}{6}e^{3t}}$ (ii) $\boldsymbol{\frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}}$ (iii) $\boldsymbol{e^{\frac{t}{2}} + 2e^{\frac{t}{3}} + 3e^{\frac{t}{3}}}$

2. *Deriving pmf from mgf.* What is the pmf of X if

$$M(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}?$$

(a) pmf A

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

(b) pmf B

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

(c) pmf C

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

3. *Expected value using mgf.* What is the expected value of

$$M(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}?$$

On the one hand, since $M(t)$ is equivalent to

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

$$E(X) = \sum_x xf(x) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{6} =$$

(i) $\mathbf{\frac{3}{3}}$ (ii) $\mathbf{\frac{4}{3}}$ (iii) $\mathbf{\frac{5}{3}}$ (iv) $\mathbf{\frac{6}{3}}$

On the other hand, since

$$M(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t},$$

the first derivative of $M(t)$ with respect to t is

$$M'(t) = \frac{d}{dt}\left[\frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}\right] = \frac{1}{2}e^t + 2 \cdot \frac{1}{3}e^{2t} + 3 \cdot \frac{1}{6}e^{3t}$$

and so evaluating this derivative at $t = 0$,

$$M'(0) = \frac{1}{2}e^0 + 2 \cdot \frac{1}{3}e^{2(0)} + 3 \cdot \frac{1}{6}e^{3(0)} = E(X) =$$

(i) $\mathbf{\frac{3}{3}}$ (ii) $\mathbf{\frac{4}{3}}$ (iii) $\mathbf{\frac{5}{3}}$ (iv) $\mathbf{\frac{6}{3}}$

which (i) **the same** (ii) **different** from before.

4. *Variance using mgf.* What is the variance of

$$M(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}?$$

On the one hand, since $M(t)$ is equivalent to

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

$$Var(X) = \sum_x (x-\mu)^2 f(x) = \left(1 - \frac{5}{3}\right)^2 \cdot \frac{1}{2} + \left(2 - \frac{5}{3}\right)^2 \cdot \frac{1}{3} + \left(3 - \frac{5}{3}\right)^2 \cdot \frac{1}{6} =$$

(i) $\mathbf{\frac{3}{9}}$ (ii) $\mathbf{\frac{4}{9}}$ (iii) $\mathbf{\frac{5}{9}}$ (iv) $\mathbf{\frac{6}{9}}$

On the other hand, since

$$M(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t},$$

the *second* derivative of $M(t)$ with respect to t is

$$M''(t) = \frac{d}{dt}\left[\frac{1}{2}e^t + 2 \cdot \frac{1}{3}e^{2t} + 3 \cdot \frac{1}{6}e^{3t}\right] = \frac{1}{2}e^t + 4 \cdot \frac{1}{3}e^{2t} + 9 \cdot \frac{1}{6}e^{3t}$$

and so evaluating this derivative at $t = 0$,

$$M''(0) = \frac{1}{2}e^0 + 2^2 \cdot \frac{1}{3}e^{2(0)} + 3^2 \cdot \frac{1}{6}e^{3(0)} = E(X^2) =$$

(i) $\mathbf{\frac{3}{3}}$ (ii) $\mathbf{\frac{4}{3}}$ (iii) $\mathbf{\frac{6}{3}}$ (iv) $\mathbf{\frac{10}{3}}$

so

$$Var(X) = E(X^2) - \mu^2 = \frac{10}{3} - \left(\frac{5}{3}\right)^2 =$$

(i) $\mathbf{\frac{3}{9}}$ (ii) $\mathbf{\frac{4}{9}}$ (iii) $\mathbf{\frac{5}{9}}$ (iv) $\mathbf{\frac{6}{9}}$

which (i) **the same** (ii) **different** from before.

5. *Binomial mgf.* With some effort, it can be shown mgf for binomial is

$$M(t) = E[e^{tX}] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \left(pe^t + q\right)^n.$$

(a) *Determine $E(X)$ using $M(t)$.*

$$E(X) = M'(0) = \left.\frac{d\left(pe^t + q\right)^n}{dt}\right]_{t=0} = \left[n\left(pe^t + q\right)^{n-1} pe^t\right]_{t=0} = n\left(pe^0 + q\right)^{n-1} pe^0 =$$

(i) $\boldsymbol{\lambda}$ (ii) $\boldsymbol{np}$ (iii) $\boldsymbol{2np}$ (iv) $\boldsymbol{npq}$.

(b) *Determine $E\left(X^2\right)$.*

$$E(X^2) = M''(0) = \left.\frac{d^2\left(pe^t + q\right)^n}{dt^2}\right]_{t=0} = \left[n(n-1)\left(pe^t + q\right)^{n-1}\left(pe^t\right)^2 + n\left(pe^t + q\right)^{n-1} pe^t\right]_{t=0}$$

$$= n(n-1)\left(pe^0 + q\right)^{n-1}\left(pe^0\right)^2 + n\left(pe^0 + q\right)^{n-1} pe^0 =$$

(i) $\boldsymbol{np(n-1)}$ (ii) $\boldsymbol{np^2(n-1)^2 + np}$ (iii) $\boldsymbol{np^2(n-1) + np}$.

(c) *Determine $Var(X)$.*

$$Var(X) = E\left(X^2\right) - E(X)^2 = (np^2(n-1) + np) - (np)^2 = np(1-p) =$$

(i) $\boldsymbol{n}$ (ii) $\boldsymbol{np}$ (iii) $\boldsymbol{2np}$ (iv) $\boldsymbol{npq}$.

6. *Identify binomial pmf with mgf.* What is the pmf of random variable X with

$$M(t) = \left(0.3e^t + 0.7\right)^{11}?$$

Since

$$\left(pe^t + q\right)^n = \left(0.3e^t + 0.7\right)^{11},$$

where $p = 0.3$, $q = 0.7$ and $n = 11$, this is a binomial distribution $b(n, p) =$

(i) $\boldsymbol{b(11, 0.3)}$ (ii) $\boldsymbol{b(0.3, 11)}$ (iii) $\boldsymbol{b(11, 0.7)}$ (iv) $\boldsymbol{b(0.7, 11)}$.

7. *Identify geometric pmf with mgf.*

(a) What is the pmf of random variable X with

$$M(t) = \frac{0.3e^t}{1 - 0.7e^t}?$$

Since, from the table above,

$$\frac{pe^t}{1 - qe^t} = \frac{0.3e^t}{1 - 0.7e^t},$$

this is a geometric distribution where

(i) $\boldsymbol{p = t, q = 0.3}$ (ii) $\boldsymbol{p = 0.7, q = 0.3}$ (iii) $\boldsymbol{p = 0.3, q = 0.7}$

(b) *Expected value.* From table above,

$$E(X) = \mu = \frac{1}{p} = \frac{1}{0.3} =$$

(i) $\frac{3}{3}$ (ii) $\frac{4}{3}$ (iii) $\frac{6}{3}$ (iv) $\frac{10}{3}$

```
library(MASS) # call up library(MASS)
fractions(1/0.3) # fractional form of 1/0.3
```

```
[1] 10/3
```

(c) *Variance.* From table above,

$$Var(X) = \sigma^2 = \frac{q}{p^2} = \frac{0.7}{0.3^2} =$$

(i) $\frac{70}{9}$ (ii) $\frac{69}{9}$ (iii) $\frac{68}{9}$ (iv) $\frac{67}{9}$

```
fractions(0.7/0.3^2) # fractional form of 0.7/0.3^2
```

```
[1] 70/9
```

8. *Deriving mgf from function of X.* If mgf of X is

$$M_X(t) = \frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}$$

then mgf of $Y = aX + b = 3X - 2$ is

$$M_Y(t) = e^{bt}M_X(at) = e^{-2t}M_X(3t) = e^{-2t}\left\{\frac{1}{2}e^{3t} + \frac{1}{3}e^{6t} + \frac{1}{6}e^{9t}\right\} =$$

(i) $\frac{1}{2}e^t + \frac{1}{3}e^{4t} + \frac{1}{6}e^{5t}$ (ii) $\frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}$ (iii) $\frac{1}{2}e^t + \frac{1}{3}e^{4t} + \frac{1}{6}e^{7t}$

9. *Deriving mgf from function of binomial X.* If mgf of X is $b(11, 0.3)$,

$$M(t) = \left(pe^t + q\right)^n = \left(0.3e^t + 0.7\right)^{11},$$

then mgf of $Y = 11 - X$ or $Y = aX + b = -X + 11$ is

$$M_Y(t) = e^{bt}M_X(at) = e^{11t}M_X(-t) = e^{11t}\left(0.3e^{-t} + 0.7\right)^{11} = \left[e^t\left(0.3e^{-t} + 0.7\right)\right]^{11} = \left[0.3 + 0.7e^t\right]^{11}$$

which is (i) $b(11, 0.3)$ (ii) $b(0.3, 11)$ (iii) $b(11, 0.7)$ (iv) $b(0.7, 11)$

10. *Binomial pgf.* With some effort, it can be shown pgf for binomial is

$$P(t) = E[t^X] = \sum_{x=0}^{n} t^x \binom{n}{x} p^x q^{n-x} = (q + pt)^n.$$

So to determine $E(X)$ using $P(t)$,

$$E(X) = P'(1) = \left.\frac{d\,(q + pt)^n}{dt}\right]_{t=1} = \left[n\,(q + pt)^{n-1}\,p\right]_{t=1} = n\,(q + p(1))^{n-1}\,p =$$

(i) λ (ii) np (iii) $2np$ (iv) npq. because $p + q = p + (1 - p) = 1$.