Pearson New International Edition

Introduction to Mathematical Statistics and Its Applications

Richard J. Larsen Morris L. Marx

Fifth Edition

ALWAYS LEARNING

PEARSON

Pearson New International Edition

Introduction to Mathematical Statistics
and Its Applications
Richard J. Larsen Morris L. Marx
Fifth Edition

PEARSON

Questions

12.9. Calculate $E(Y^3)$ for a random variable whose moment-generating function is $M_Y(t) = e^{t^2/2}$.

12.10. Find $E(Y^4)$ if Y is an exponential random variable with $f_Y(y) = \lambda e^{-\lambda y}$, $y > 0$.

12.11. The form of the moment-generating function for a normal random variable is $M_Y(t) = e^{at+b^2t^2/2}$ (recall Example 12.4). Differentiate $M_Y(t)$ to verify that $a = E(Y)$ and $b^2 = \text{Var}(Y)$.

12.12. What is $E(Y^4)$ if the random variable Y has moment-generating function $M_Y(t) = (1 - \alpha t)^{-k}$?

12.13. Find $E(Y^2)$ if the moment-generating function for Y is given by $M_Y(t) = e^{-t+4t^2}$. Use Example 12.4 to find $E(Y^2)$ without taking any derivatives. (*Hint*: Recall Theorem 6.1.)

12.14. Find an expression for $E(Y^k)$ if $M_Y(t) = (1 - t/\lambda)^{-r}$, where λ is any positive real number and r is a positive integer.

12.15. Use $M_Y(t)$ to find the expected value of the uniform random variable described in Question 12.1.

12.16. Find the variance of Y if $M_Y(t) = e^{2t}/(1 - t^2)$.

Using Moment-Generating Functions to Identify Pdfs

Finding moments is not the only application of moment-generating functions. They are also used to identify the pdf of *sums* of random variables—that is, finding $f_W(w)$, where $W = W_1 + W_2 + \cdots + W_n$. Their assistance in the latter is particularly important for two reasons: (1) Many statistical procedures are defined in terms of sums, and (2) alternative methods for deriving $f_{W_1+W_2+\cdots+W_n}(w)$ are extremely cumbersome.

The next two theorems give the background results necessary for deriving $f_W(w)$. Theorem 12.2 states a key uniqueness property of moment-generating functions: If W_1 and W_2 are random variables with the same mgfs, they must necessarily have the same pdfs. In practice, applications of Theorem 12.2 typically rely on one or both of the algebraic properties cited in Theorem 12.3.

Theorem 12.2 *Suppose that W_1 and W_2 are random variables for which $M_{W_1}(t) = M_{W_2}(t)$ for some interval of t's containing* 0. *Then $f_{W_1}(w) = f_{W_2}(w)$.*

Proof See (95). □

Theorem 12.3

a. *Let W be a random variable with moment-generating function $M_W(t)$. Let $V = aW + b$. Then*

$$M_V(t) = e^{bt} M_W(at)$$

b. *Let $W_1, W_2, \ldots, W_n$ be independent random variables with moment-generating functions $M_{W_1}(t), M_{W_2}(t), \ldots,$ and $M_{W_n}(t)$, respectively. Let $W = W_1 + W_2 + \cdots + W_n$. Then*

$$M_W(t) = M_{W_1}(t) \cdot M_{W_2}(t) \cdots M_{W_n}(t)$$

Proof The proof is left as an exercise. □

Example 12.10 Suppose that X_1 and X_2 are two independent Poisson random variables with parameters λ_1 and λ_2, respectively. That is,

$$p_{X_1}(k) = P(X_1 = k) = \frac{e^{-\lambda_1}\lambda_1 k}{k!}, \quad k = 0, 1, 2, \ldots$$

and

$$p_{X_2}(k) = P(X_2 = k) = \frac{e^{-\lambda_2}\lambda_2 k}{k!}, \quad k = 0, 1, 2 \ldots$$

Let $X = X_1 + X_2$. What is the pdf for X?

According to Example 12.9, the moment-generating functions for X_1 and X_2 are

$$M_{X_1}(t) = e^{-\lambda_1 + \lambda_1 e^t}$$

and

$$M_{X_2}(t) = e^{-\lambda_2 + \lambda_2 e^t}$$

Moreover, if $X = X_1 + X_2$, then by part (b) of Theorem 12.3,

$$\begin{aligned} M_X(t) &= M_{X_1}(t) \cdot M_{X_2}(t) \\ &= e^{-\lambda_1 + \lambda_1 e^t} \cdot e^{-\lambda_2 + \lambda_2 e^t} \\ &= e^{-(\lambda_1 + \lambda_2) + (\lambda_1 + \lambda_2)e^t} \end{aligned} \tag{12.7}$$

But, by inspection, Equation 12.7 is the moment-generating function that a Poisson random variable with $\lambda = \lambda_1 + \lambda_2$ would have. It follows, then, by Theorem 12.2 that

$$p_X(k) = \frac{e^{-(\lambda_1 + \lambda_2)}(\lambda_1 + \lambda_2)^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Comment The Poisson random variable reproduces itself in the sense that the sum of independent Poissons is also a Poisson. A similar property holds for independent normal random variables (see Question 12.19) and, under certain conditions, for independent binomial random variables (recall Example 8.2). ■

Example 12.11 We saw in Example 12.4 that a normal random variable, Y, with mean μ and variance σ^2 has pdf

$$f_Y(y) = \left(1/\sqrt{2\pi}\sigma\right) \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right], \quad -\infty < y < \infty$$

and mgf

$$M_Y(t) = e^{\mu t + \sigma^2 t^2/2}$$

By definition, a *standard normal random variable* is a normal random variable for which $\mu = 0$ and $\sigma = 1$. Denoted Z, the pdf and mgf for a standard normal random variable are $f_Z(z) = (1/\sqrt{2\pi})e^{-z^2/2}$, $-\infty < z < \infty$, and $M_Z(t) = e^{t^2/2}$, respectively. Show that the ratio

$$\frac{Y-\mu}{\sigma}$$

is a standard normal random variable, Z.

Write $\frac{Y-\mu}{\sigma}$ as $\frac{1}{\sigma}Y - \frac{\mu}{\sigma}$. By part (a) of Theorem 12.3,

$$\begin{aligned} M_{(Y-\mu)/\sigma}(t) &= e^{-\mu t/\sigma} M_Y\left(\frac{t}{\sigma}\right) \\ &= e^{-\mu t/\sigma} e^{[\mu t/\sigma + \sigma^2 (t/\sigma)^2/2]} \\ &= e^{t^2/2} \end{aligned}$$

But $M_Z(t) = e^{t^2/2}$ so it follows from Theorem 12.2 that the pdf for $\frac{Y-\mu}{\sigma}$ is the same as the pdf for $f_z(z)$. (We call $\frac{Y-\mu}{\sigma}$ a *Z transformation*.) ■

Questions

12.17. Use Theorem 12.3(a) and Question 12.8 to find the moment-generating function of the random variable Y, where $f_Y(y) = \lambda y e^{-\lambda y}$, $y \geq 0$.

12.18. Let Y_1, Y_2, and Y_3 be independent random variables, each having the pdf of Question 12.17. Use Theorem 12.3(b) to find the moment-generating function of $Y_1 + Y_2 + Y_3$. Compare your answer to the moment-generating function in Question 12.14.

12.19. Use Theorems 12.2 and 12.3 to determine which of the following statements is true:

(a) The sum of two independent Poisson random variables has a Poisson distribution.
(b) The sum of two independent exponential random variables has an exponential distribution.
(c) The sum of two independent normal random variables has a normal distribution.

12.20. Calculate $P(X \leq 2)$ if $M_X(t) = \left(\frac{1}{4} + \frac{3}{4}e^t\right)^5$.

12.21. Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample of size n from a normal distribution with mean μ and standard deviation σ. Use moment-generating functions to deduce the pdf of $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

12.22. Suppose the moment-generating function for a random variable W is given by

$$M_W(t) = e^{-3+3e^t} \cdot \left(\frac{2}{3} + \frac{1}{3}e^t\right)^4$$

Calculate $P(W \leq 1)$. (*Hint:* Write W as a sum.)

12.23. Suppose that X is a Poisson random variable, where $p_X(k) = e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \ldots$.

(a) Does the random variable $W = 3X$ have a Poisson distribution?
(b) Does the random variable $W = 3X + 1$ have a Poisson distribution?

12.24. Suppose that Y is a normal variable, where $f_Y(y) = (1/\sqrt{2\pi}\sigma)\exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$, $-\infty < y < \infty$.

(a) Does the random variable $W = 3Y$ have a normal distribution?
(b) Does the random variable $W = 3Y + 1$ have a normal distribution?

13 Taking a Second Look at Statistics (Interpreting Means)

One of the most important ideas is the notion of the *expected value* (or *mean*) of a random variable. Defined in Section 5 as a number that reflects the "center" of a pdf, the expected value (μ) was originally introduced for the benefit of gamblers. It spoke directly to one of their most fundamental questions—How much will I win or lose, *on the average,* if I play a certain game? (Actually, the real question they probably had in mind was "How much are *you* going to *lose,* on the average?") Despite having had such a selfish, materialistic, gambling-oriented *raison d'etre,* the expected value was quickly embraced by (respectable) scientists and researchers of all persuasions as a preeminently useful descriptor of a distribution. Today, it would not be an exaggeration to claim that the majority of *all* statistical analyses focus on either (1) the expected value of a single random variable or (2) comparing the expected values of two or more random variables.

In the lingo of applied statistics, there are actually two fundamentally different types of "means"—*population means* and *sample means*. The term "population mean" is a synonym for what mathematical statisticians would call an expected value—that is, a population mean (μ) is a weighted average of the possible values associated with a theoretical probability model, either $p_X(k)$ or $f_Y(y)$, depending on whether the underlying random variable is discrete or continuous. A *sample mean* is the arithmetic average of a set of measurements. If, for example, n observations—$y_1, y_2, \ldots, y_n$—are taken on a continuous random variable Y, the sample mean is denoted $\bar{y}$, where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Conceptually, sample means are *estimates* of population means, where the "quality" of the estimation is a function of (1) the sample size and (2) the standard deviation (σ) associated with the individual measurements. Intuitively, as the sample size gets larger and/or the standard deviation gets smaller, the approximation will tend to get better.

Interpreting means (either $\bar{y}$ or μ) is not always easy. To be sure, what they imply *in principle* is clear enough—both $\bar{y}$ and μ are measuring the centers of their respective distributions. Still, many a wrong conclusion can be traced directly to researchers misunderstanding the value of a mean. Why? Because the distributions that $\bar{y}$ and/or μ are *actually* representing may be dramatically different from the distributions we *think* they are representing.

An interesting case in point arises in connection with SAT scores. Each fall the average SATs earned by students in each of the fifty states and the District of Columbia are released by the Educational Testing Service (ETS). With "accountability" being one of the new paradigms and buzzwords associated with K–12 education, SAT scores have become highly politicized. At the national level, Democrats and Republicans each campaign on their own versions of education reform, fueled in no small measure by scores on standardized exams, SATs included; at the state level, legislatures often modify education budgets in response to how well or how poorly their students performed the year before. Does it make sense, though, to use SAT averages to characterize the quality of a state's education system? Absolutely not! Averages of this sort refer to very different distributions from state to state. Any attempt to interpret them at face value will necessarily be misleading.

One such state-by-state SAT comparison that appeared in the mid-90s is reproduced in Table 13.1. Notice that Tennessee's entry is 1023, which is the tenth highest average listed. Does it follow that Tennessee's educational system is among the best in the nation? Probably not. Most independent assessments of K–12 education rank Tennessee's schools among the weakest in the nation, not among the best. If those opinions are accurate, why do Tennessee's students do so well on the SAT?

The answer to that question lies in the academic profiles of the students who take the SAT in Tennessee. Most college-bound students in that state apply exlusively to schools in the South and the Midwest, where admissions are based on the ACT, not the SAT. The SAT is primarily used by private schools, where admissions tend to be more competitive. As a result, the students in Tennessee who take the SAT are not representative of the entire population of students in that state. A disproportionate number are exceptionally strong academically, those being the students who feel that they have the ability to be

Table 13.1

State	Average SAT Score	State	Average SAT Score
AK	911	MT	986
AL	1011	NE	1025
AZ	939	NV	913
AR	935	NH	924
CA	895	NJ	893
CO	969	NM	1003
CT	898	NY	888
DE	892	NC	860
DC	849	ND	1056
FL	879	OH	966
GA	844	OK	1019
HI	881	OR	927
ID	969	PA	879
IL	1024	RI	882
IN	876	SC	838
IA	1080	SD	1031
KS	1044	TN	1023
KY	997	TX	886
LA	1011	UT	1067
ME	883	VT	899
MD	908	VA	893
MA	901	WA	922
MI	1009	WV	921
MN	1057	WI	1044
MS	1013	WY	980
MO	1017		

competitive at Ivy League–type schools. The number 1023, then, is the average of *something* (in this case, an elite subset of all Tennessee students), but it does not correspond to the center of the SAT distribution for *all* Tennessee students.

The moral here is that analyzing data effectively requires that we look beyond the obvious. What we have learned in this chapter about random variables and probability distributions and expected values will be helpful only if we take the time to learn about the context and the idiosyncrasies of the phenomenon being studied. To do otherwise is likely to lead to conclusions that are, at best, superficial and, at worst, incorrect.

Appendix A.1 Minitab Applications

Numerous software packages are available for performing a variety of probability and statistical calculations. Among the first to be developed, and one that continues to be very popular, is Minitab. Here we include a short discussion of Minitab solutions to some of the problems that were discussed in the chapter. What other software packages can do and the ways their outputs are formatted are likely to be quite similar.

Contained in Minitab are subroutines that can do some of the more important pdf and cdf computations described in Sections 3 and 4. In the case of binomial random variables, for instance, the statements

```
MTB  > pdf k;
SUBC > binomial n p.
```

and

```
MTB  > cdf k;
SUBC > binomial n p.
```

will calculate $\binom{n}{k}p^k(1-p)^{n-k}$ and $\sum_{r=0}^{k}\binom{n}{r}p^r(1-p)^{n-r}$, respectively. Figure A.1.1 shows the Minitab program for doing the cdf calculation $[= P(X \leq 15)]$ asked for in part (a) of Example 2.2.

The commands `pdf k` and `cdf k` can be run on many of the probability models most likely to be encountered in real-world problems. Those on the list that we have already seen are the binomial, Poisson, normal, uniform, and exponential distributions.

Figure A.1.1

```
MTB > cdf 15;
SUBC > binomial 30 0.60.
```

Cumulative Distribution Function

```
Binomial with n = 30 and p = 0.600000
      x    P(X <= x)
    15.00   0.1754
```

For discrete random variables, the cdf can be printed out in its entirety (that is, for every integer) by deleting the argument k and using the command `MTB < cdf;`. Typical is the output in Figure A.1.2, corresponding to the cdf for a binomial random variable with $n = 4$ and $p = \frac{1}{6}$.

Figure A.1.2

```
MTB > cdf;
SUBC > binomial 4 0.167.
```

Cumulative Distribution Function

```
Binomial with n = 4 and p =0.167000
        x    P( X <= x)
        0       0.4815
        1       0.8676
        2       0.9837
        3       0.9992
        4       1.0000
```

Figure A.1.3

```
MTB > invcdf 0.60;
SUBC > exponential 1.
```

Inverse Cumulative Distribution Function

```
Exponential with mean = 1.00000
   P(X <= x)  x
    0.6000   0.9163
```

Also available is an *inverse cdf* command, which in the case of a continuous random variable Y and a specified probability p identifies the value y having the