21 Explain why the surface area is infinite when $y = 1/x$ is rotated around the x axis ($1 \leqslant x < \infty$). But the volume of "Gabriel's horn" is ________. It can't hold enough paint to paint its surface.

22 A disk of radius 1″ can be covered by four strips of tape (width $\frac{1}{2}$″). If the strips are not parallel, prove that they can't cover the disk. **Hint**: Change to a unit sphere sliced by planes $\frac{1}{2}$″ apart. Problem 14 gives surface area π for each slice.

23 A watermelon (maybe a football) is the result of rotating half of the ellipse $x = \sqrt{2}\cos t$, $y = \sin t$ (which means $x^2 + 2y^2 = 2$). Find the surface area, parametrically or not.

24 Estimate the surface area of an egg.

8.4 Probability and Calculus

Discrete probability usually involves careful counting. Not many samples are taken and not many experiments are made. There is a list of possible outcomes, and a known probability for each outcome. But probabilities go far beyond red cards and black cards. The real questions are much more practical:

1. How often will too many passengers arrive for a flight?
2. How many random errors do you make on a quiz?
3. What is the chance of exactly one winner in a big lottery?

Those are important questions and we will set up models to answer them.

There is another point. Discrete models do not involve calculus. The number of errors or bumped passengers or lottery winners is a small whole number. ***Calculus enters for continuous probability.*** Instead of results that exactly equal 1 or 2 or 3, calculus deals with results that fall in a range of numbers. Continuous probability comes up in at least two ways:

(A) An experiment is repeated many times and we take *averages*.
(B) The outcome lies anywhere in an *interval* of numbers.

In the continuous case, the probability p_n of hitting a particular value $x = n$ becomes zero. Instead we have a ***probability density*** $p(x)$—which is a key idea. *The chance that a random X falls between a and b is found by integrating the density $p(x)$:*

$$\text{Prob}\{a \leqslant X \leqslant b\} = \int_a^b p(x)\,dx. \tag{1}$$

Roughly speaking, $p(x)\,dx$ is the chance of falling between x and $x + dx$. Certainly $p(x) \geqslant 0$. If a and b are the extreme limits $-\infty$ and ∞, including all possible outcomes, the probability is necessarily one:

$$\text{Prob}\{-\infty < X < +\infty\} = \int_{-\infty}^{\infty} p(x)\,dx = 1. \tag{2}$$

This is a case where infinite limits of integration are natural and unavoidable. In studying probability they create no difficulty—areas out to infinity are often easier.

Here are typical questions involving continuous probability and calculus:

4. How conclusive is a 53%–47% poll of 2500 voters?
5. Are 16 random football players safe on an elevator with capacity 3600 pounds?
6. How long before your car is in an accident?

It is not so traditional for a calculus course to study these questions. They need extra thought, beyond computing integrals (so this section is harder than average). But probability is more important than some traditional topics, and also more interesting.

Drug testing and gene identification and market research are major applications. Comparing Questions 1–3 with **4–6** brings out the relation of **discrete** to **continuous**—the differences between them, and the parallels.

It would be impossible to give here a full treatment of probability theory. I believe you will see the point (and the use of calculus) from our examples. Frank Morgan's lectures have been a valuable guide.

DISCRETE RANDOM VARIABLES

A **discrete** random variable X has a list of possible values. For two dice the outcomes are $X = 2, 3, \ldots, 12$. For coin tosses (see below), the list is infinite: $X = 1, 2, 3, \ldots$.

A ***continuous*** variable lies in an interval $a \leqslant X \leqslant b$.

EXAMPLE 1 Toss a fair coin until heads come up. The outcome X is the ***number of tosses***. The value of X is 1 or 2 or 3 or ..., and the probability is $\frac{1}{2}$ that $X = 1$ (heads on the first toss). The probability of tails then heads is $p_2 = \frac{1}{4}$. The probability that $X = n$ is $p_n = (\frac{1}{2})^n$—this is the chance of $n - 1$ tails followed by heads. ***The sum of all probabilities is necessarily*** 1:

$$p_1 + p_2 + p_3 + \cdots = \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \cdots = 1.$$

EXAMPLE 2 Suppose a student (not you) makes an average of 2 unforced errors per hour exam. The number of actual errors on the next exam is $X = 0$ or 1 or 2 or .... A reasonable model for the probability of n errors—when they are random and independent—is the ***Poisson model*** (pronounced Pwason):

$$p_n = \textit{probability of n errors} = \frac{2^n}{n!} e^{-2}.$$

The probabilities of no errors, one error, and two errors are p_0, p_1, and p_2:

$$p_0 = \frac{2^0}{0!} e^{-2} = \frac{1}{1} e^{-2} \approx .135 \qquad p_1 = \frac{2^1}{1!} e^{-2} \approx .27 \qquad p_2 = \frac{2^2}{2!} e^{-2} \approx .27.$$

The probability of more than two errors is $1 - .135 - .27 - .27 = .325$.

This Poisson model can be derived theoretically or tested experimentally. The total probability is again 1, from the infinite series (Section 6.6) for e^2:

$$p_0 + p_1 + p_2 + \cdots = \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \cdots\right) e^{-2} = e^2 e^{-2} = 1. \qquad (3)$$

EXAMPLE 3 Suppose on average 3 out of 100 passengers with reservations don't show up for a flight. If the plane holds 98 passengers, ***what is the probability that someone will be bumped?***

If the passengers come independently to the airport, use the Poisson model with 2 changed to 3. X is the number of no-shows, and $X = n$ happens with probability p_n:

$$p_n = \frac{3^n}{n!} e^{-3} \qquad p_0 = \frac{3^0}{0!} e^{-3} = e^{-3} \qquad p_1 = \frac{3^1}{1!} e^{-3} = 3e^{-3}.$$

There are 98 seats and 100 reservations. Someone is bumped if $X = 0$ or $X = 1$:

$$\text{chance of bumping} = p_0 + p_1 = e^{-3} + 3e^{-3} \approx 4/20.$$

We will soon define the ***average*** or ***expected value*** or ***mean*** of X—this model has $\mu = 3$.

CONTINUOUS RANDOM VARIABLES

If X is the lifetime of a VCR, all numbers $X \geqslant 0$ are possible. If X is a score on the SAT, then $200 \leqslant X \leqslant 800$. If X is the fraction of computer owners in a poll of 600 people, X is between 0 and 1. You may object that the SAT score is a whole number and the fraction of computer owners must be 0 or 1/600 or 2/600 or .... But it is completely impractical to work with 601 discrete possibilities. Instead we take X to be a ***continuous random variable***, falling *anywhere* in the range $X \geqslant 0$ or [200, 800] or $0 \leqslant X \leqslant 1$. Of course the various values of X are not equally probable.

EXAMPLE 4 The average lifetime of a VCR is 4 years. A reasonable model for breakdown time is an ***exponential random variable***. Its probability density is

$$p(x) = \tfrac{1}{4}e^{-x/4} \quad \text{for} \quad 0 \leqslant x < \infty.$$

The probability that the VCR will eventually break is 1:

$$\int_0^\infty \tfrac{1}{4}e^{-x/4}\,dx = \Big[-e^{-x/4}\Big]_0^\infty = 0 - (-1) = 1. \tag{4}$$

The probability of breakdown within 12 years (X from 0 to 12) is .95:

$$\int_0^{12} \tfrac{1}{4}e^{-x/4}\,dx = \Big[-e^{-x/4}\Big]_0^{12} = -e^{-3} + 1 \approx .95. \tag{5}$$

An exponential distribution has $p(x) = ae^{-ax}$. Its integral from 0 to x is $F(x) = 1 - e^{-ax}$. Figure 8.11 is the graph for $a = 1$. It shows the area up to $x = 1$.

To repeat: ***The probability that*** $a \leqslant X \leqslant b$ ***is the integral of*** $p(x)$ ***from*** a ***to*** b.

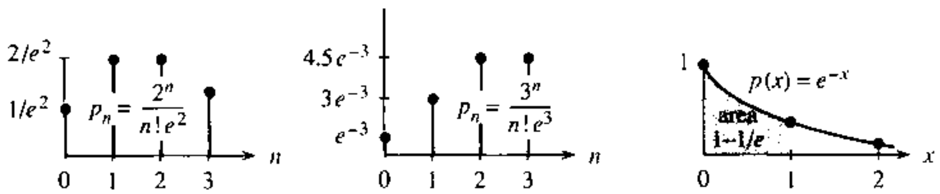

Fig. 8.11 Probabilities add to $\Sigma p_n = 1$. Continuous density integrates to $\int p(x)\,dx = 1$.

EXAMPLE 5 We now define the most important density function. Suppose the average SAT score is 500, and the *standard deviation* (defined below—it measures the spread around the average) is 200. Then the ***normal distribution*** of grades has

$$p(x) = \frac{1}{200\sqrt{2\pi}}e^{-(x-500)^2/2(200)^2} \quad \text{for} \quad -\infty < x < \infty.$$

This is the normal (or Gaussian) distribution with mean 500 and standard deviation 200. The graph of $p(x)$ is the famous ***bell-shaped curve*** in Figure 8.12.

A new objection is possible. The actual scores are between 200 and 800, while the density $p(x)$ extends all the way from $-\infty$ to ∞. I think the Educational Testing Service counts all scores over 800 as 800. The fraction of such scores is pretty small—in fact the normal distribution gives

$$\text{Prob}\{X \geqslant 800\} = \int_{800}^\infty \frac{1}{200\sqrt{2\pi}}e^{-(x-500)^2/2(200)^2}\,dx \approx .0013. \tag{6}$$

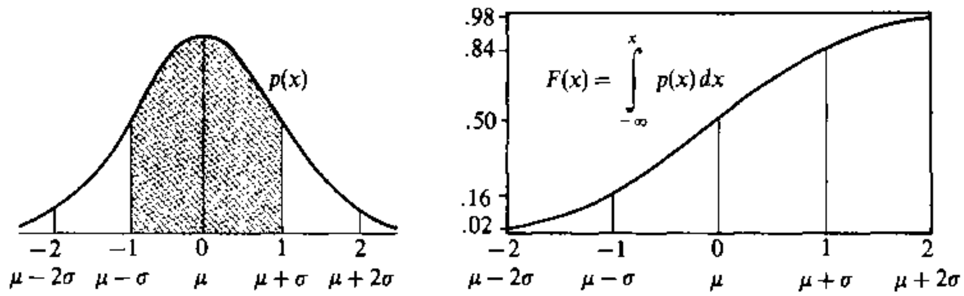

Fig. 8.12 The normal distribution (bell-shaped curve) and its cumulative density $F(x)$.

Regrettably, e^{-x^2} has no elementary antiderivative. We need numerical integration. But there is nothing the matter with that! The integral is called the "*error function*," and special tables give its value to great accuracy. The integral of $e^{-x^2/2}$ from $-\infty$ to ∞ is exactly $\sqrt{2\pi}$. Then division by $\sqrt{2\pi}$ keeps $\int p(x)\,dx = 1$.

Notice that the normal distribution involves *two parameters*. They are the mean value (in this case $\mu = 500$) and the standard deviation (in this case $\sigma = 200$). Those numbers *mu* and *sigma* are often given the "normalized" values $\mu = 0$ and $\sigma = 1$:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad \text{becomes} \quad p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

The bell-shaped graph of p is symmetric around the middle point $x = \mu$. The width of the graph is governed by the second parameter σ—which stretches the x axis and shrinks the y axis (leaving total area equal to 1). The axes are labeled to show the standard case $\mu = 0$, $\sigma = 1$ and also the graph for any other μ and σ.

We now give a name to the integral of $p(x)$. The limits will be $-\infty$ and x, so the integral $F(x)$ measures the ***probability that a random sample is below*** x:

$$\text{Prob}\{X \leqslant x\} = \int_{-\infty}^{x} p(x)\,dx = \textbf{\textit{cumulative density function}}\ F(x). \qquad (7)$$

$F(x)$ accumulates the probabilities given by $p(x)$, so $dF/dx = p(x)$. The total probability is $F(\infty) = 1$. This integral from $-\infty$ to ∞ covers all outcomes.

Figure 8.12b shows the integral of the bell-shaped normal distribution. The middle point $x = \mu$ has $F = \frac{1}{2}$. By symmetry there is a 50-50 chance of an outcome below the mean. The cumulative density $F(x)$ is near .16 at $\mu - \sigma$ and near .84 at $\mu + \sigma$. The chance of falling in between is $.84 - .16 = .68$. Thus 68% of the outcomes are less than one deviation σ away from the center μ.

Moving out to $\mu - 2\sigma$ and $\mu + 2\sigma$, 95% of the area is in between. ***With 95% confidence X is less than two deviations from the mean.*** Only one sample in 20 is further out (less than one in 40 on each side).

Note that $\sigma = 200$ is not the precise value for the SAT!

MEAN, VARIANCE, AND STANDARD DEVIATION

In Example 1, X was the number of coin tosses until the appearance of heads. The probabilities were $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$, $p_3 = \frac{1}{8}$, ... *What is the average number of tosses?* We now find the "mean" μ of any distribution $p(x)$—not only the normal distribution, where symmetry guarantees that the built-in number μ is the mean.

To find μ, ***multiply outcomes by probabilities and add***:

$$\mu = \textbf{\textit{mean}} = \sum n p_n = 1(p_1) + 2(p_2) + 3(p_3) + \cdots. \qquad (8)$$

The average number of tosses is $1(\frac{1}{2})+2(\frac{1}{4})+3(\frac{1}{8})+\cdots$. This series adds up (in Section 10.1) to $\mu=2$. Please do the experiment 10 times. I am almost certain that the average will be near 2.

When the average is $\lambda=2$ quiz errors or $\lambda=3$ no-shows, the Poisson probabilities are $p_n=\lambda^n e^{-\lambda}/n!$ Check that the formula $\mu=\Sigma\, np_n$ does give λ as the mean:

$$\left[1\frac{\lambda}{1!}+2\frac{\lambda^2}{2!}+3\frac{\lambda^3}{3!}+\cdots\right]e^{-\lambda}=\lambda\left[1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\cdots\right]e^{-\lambda}=\lambda e^{\lambda}e^{-\lambda}=\lambda.$$

For continuous probability, the sum $\mu=\Sigma\, np_n$ changes to $\mu=\int xp(x)\,dx$. We multiply outcome x by probability $p(x)$ and integrate. In the VCR model, integration by parts gives a mean breakdown time of $\mu=4$ years:

$$\int x\,p(x)\,dx=\int_0^\infty x(\tfrac{1}{4}e^{-x/4})\,dx=\Big[-xe^{-x/4}-4e^{-x/4}\Big]_0^\infty=4. \qquad (9)$$

Together with the mean we introduce the ***variance***. It is always written σ^2, and in the normal distribution that measured the "width" of the curve. When σ^2 was 200^2, SAT scores spread out pretty far. If the testing service changed to $\sigma^2=1^2$, the scores would be a disaster. 95% of them would be within ± 2 of the mean. When a teacher announces an average grade of 72, the variance should also be announced—if it is big then those with 60 can relax. At least they have company.

8E The mean μ is the expected value of X. The variance σ^2 is the expected value of $(X-\text{mean})^2=(X-\mu)^2$. Multiply outcome times probability and add:

$\mu=\sum np_n$	$\sigma^2=\sum(n-\mu)^2p_n$	(discrete)
$\mu=\int_{-\infty}^{\infty}xp(x)\,dx$	$\sigma^2=\int_{-\infty}^{\infty}(x-\mu)^2p(x)\,dx$	(continuous)

The ***standard deviation*** (written σ) is the square root of σ^2.

EXAMPLE 6 (Yes-no poll, one person asked) The probabilities are p and $1-p$.

A fraction $p=\frac{1}{3}$ of the population thinks *yes*, the remaining fraction $1-p=\frac{2}{3}$ thinks *no*. Suppose we only ask one person. If $X=1$ for yes and $X=0$ for no, the expected value of X is $\mu=p=\frac{1}{3}$. The variance is $\sigma^2=p(1-p)=\frac{2}{9}$:

$$\mu=0\left(\frac{2}{3}\right)+1\left(\frac{1}{3}\right)=\frac{1}{3}\quad\text{and}\quad\sigma^2=\left(0-\frac{1}{3}\right)^2\left(\frac{2}{3}\right)+\left(1-\frac{1}{3}\right)^2\left(\frac{1}{3}\right)=\frac{2}{9}.$$

The standard deviation is $\sigma=\sqrt{2/9}$. When the fraction p is near one or near zero, the spread is smaller—and one person is more likely to give the right answer for everybody. The maximum of $\sigma^2=p(1-p)$ is at $p=\frac{1}{2}$, where $\sigma=\frac{1}{2}$.

The table shows μ and σ^2 for important probability distributions.

Model	*Mean*	*Variance*	*Application*
$p_1=p,\ p_0=1-p$	p	$p(1-p)$	yes-no
Poisson $p_n=\lambda^n e^{-\lambda}/n!$	λ	λ	random occurrence
Exponential $p(x)=ae^{-ax}$	$1/a$	$1/a^2$	waiting time
Normal $p(x)=\dfrac{1}{\sqrt{2\pi}\,\sigma}e^{-(x-\mu)^2/2\sigma^2}$	μ	σ^2	distribution around mean

THE LAW OF AVERAGES AND THE CENTRAL LIMIT THEOREM

We come to the center of probability theory (without intending to give proofs). The key idea is to repeat an experiment many times—poll many voters, or toss many dice, or play considerable poker. Each independent experiment produces an outcome X, and the average from N experiments is $\bar{X}$. It is called "X *bar*":

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_N}{N} = \text{average outcome.}$$

All we know about $p(x)$ is its mean μ and variance σ^2. It is amazing how much information that gives about the average $\bar{X}$:

> **8F** ***Law of Averages***: $\bar{X}$ is almost sure to approach μ as $N \to \infty$.
> ***Central Limit Theorem***: The probability density $p_N(x)$ for $\bar{X}$ approaches a normal distribution with the same mean μ and with variance σ^2/N.

No matter what the probabilities for X, the probabilities for $\bar{X}$ move toward the normal bell-shaped curve. The standard deviation is close to $\sigma/\sqrt{N}$ when the experiment is repeated N times. In the Law of Averages, "almost sure" means that the chance of $\bar{X}$ *not* approaching μ is zero. It can happen, but it won't.

Remark 1 The Boston Globe doesn't understand the Law of Averages. I quote from September 1988: "What would happen if a giant Red Sox slump arrived? What would happen if the fabled Law of Averages came into play, reversing all those can't miss decisions during the winning streak?" They think the Law of Averages evens everything up, favoring heads after a series of tails. See Problem 20.

EXAMPLE 7 *Yes-no poll of $N = 2500$ voters. **Is a** 53%–47% **outcome conclusive**?*

The fraction p of "yes" voters in the whole population is *not known*. That is the reason for the poll. The deviation $\sigma = \sqrt{p(1-p)}$ is also not known, but for one voter this is never more than $\frac{1}{2}$ (when $p = \frac{1}{2}$). Therefore $\sigma/\sqrt{N}$ for 2500 voters is no larger than $\frac{1}{2}/\sqrt{2500}$, which is 1%.

The result of the poll was $\bar{X} = 53\%$. With 95% confidence, this sample is within two standard deviations (here 2%) of its mean. Therefore with 95% confidence, ***the unknown mean $\mu = p$ of the whole population is between*** 51% ***and*** 55%. This poll is conclusive.

If the true mean had been $p = 50\%$, the poll would have had only a .0013 chance of reaching 53%. The error margin on each side of a poll is amazingly simple; it is always $1/\sqrt{N}$.

Remark 2 The New York Times has better mathematicians than the Globe. Two days after Bush defeated Dukakis, their poll of $N = 11{,}645$ voters was printed with the following explanation. "In theory, in 19 cases out of 20 [*there is* 95%] the results should differ by no more than one percentage point [*there is* $1/\sqrt{N}$] from what would have been obtained by seeking out all voters in the United States."

EXAMPLE 8 Football players at Caltech (if any) have average weight $\mu = 210$ pounds and standard deviation $\sigma = 30$ pounds. Are $N = 16$ players safe on an elevator with capacity 3600 pounds? 16 times 210 is 3360.

The average weight $\bar{X}$ is approximately a normal random variable with $\bar{\mu} = 210$ and $\bar{\sigma} = 30/\sqrt{N} = 30/4$. There is only a 2% chance that $\bar{X}$ is above $\bar{\mu} + 2\bar{\sigma} = 225$ (see Figure 8.12b—weights below the mean are no problem on an elevator). Since 16 times 225 is 3600, a statistician would have 98% confidence that the elevator is safe. This is an example where 98% is not good enough—I wouldn't get on.

EXAMPLE 9 (The famous Weldon Dice) Weldon threw 12 dice 26,306 times and counted the 5's and 6's. They came up in 33.77% of the 315,672 separate rolls. Thus $\bar{X} = .3377$ instead of the expected fraction $p = \frac{1}{3}$ of 5's and 6's. Were the dice fair?

The variance in each roll is $\sigma^2 = p(1-p) = 2/9$. The standard deviation of $\bar{X}$ is $\bar{\sigma} = \sigma/\sqrt{N} = \sqrt{2/9}/\sqrt{315672} \approx .00084$. For fair dice, there is a 95% chance that $\bar{X}$ will differ from $\frac{1}{3}$ by less than $2\bar{\sigma}$. (For Poisson probabilities that is false. Here $\bar{X}$ is *normal*.) But .3377 differs from .3333 by more than $5\bar{\sigma}$. The chance of falling 5 standard deviations away from the mean is only about 1 in 10,000.†

So the dice were unfair. The faces with 5 or 6 indentations were *lighter* than the others, and a little more likely to come up. Modern dice are made to compensate for that, but Weldon never tried again.

8.4 EXERCISES

Read-through questions

Discrete probability uses counting, ___a___ probability uses calculus. The function $p(x)$ is the probability ___b___. The chance that a random variable falls between a and b is ___c___. The total probability is $\int_{-\infty}^{\infty} p(x)\,dx =$ ___d___. In the discrete case $\Sigma\, p_n =$ ___e___. The mean (or expected value) is $\mu = \int$ ___f___ in the continuous case and $\mu = \Sigma\, np_n$ in the ___g___.

The Poisson distribution with mean λ has $p_n =$ ___h___. The sum $\Sigma\, p_n = 1$ comes from the ___i___ series. The exponential distribution has $p(x) = e^{-x}$ or $2e^{-2x}$ or ___j___. The standard Gaussian (or ___k___) distribution has $\sqrt{2\pi}\,p(x) = e^{-x^2/2}$. Its graph is the well-known ___l___ curve. The chance that the variable falls below x is $F(x) =$ ___m___. F is the ___n___ density function. The difference $F(x+dx) - F(x)$ is about ___o___, which is the chance that X is between x and $x + dx$.

The *variance*, which measures the spread around μ, is $\sigma^2 = \int$ ___p___ in the continuous case and $\sigma^2 = \Sigma$ ___q___ in the discrete case. Its square root σ is the ___r___. The normal distribution has $p(x) =$ ___s___. If $\bar{X}$ is the ___t___ of N samples from any population with mean μ and variance σ^2, the Law of Averages says that $\bar{X}$ will approach ___u___. The Central Limit Theorem says that the distribution for $\bar{X}$ approaches ___v___. Its mean is ___w___ and its variance is ___x___.

In a yes-no poll when the voters are 50-50, the mean for one voter is $\mu = 0(\frac{1}{2}) + 1(\frac{1}{2}) =$ ___y___. The variance is $(0-\mu)^2 p_0 + (1-\mu)^2 p_1 =$ ___z___. For a poll with $N = 100$, $\bar{\sigma}$ is ___A___. There is a 95% chance that $\bar{X}$ (the fraction saying yes) will be between ___B___ and ___C___.

1 If $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$, $p_3 = \frac{1}{8}$, ..., what is the probability of an outcome $X < 4$? What are the probabilities of $X = 4$ and $X > 4$?

2 With the same $p_n = (\frac{1}{2})^n$, what is the probability that X is odd? Why is $p_n = (\frac{1}{3})^n$ an impossible set of probabilities? What multiple $c(\frac{1}{3})^n$ is possible?

3 Why is $p(x) = e^{-2x}$ not an acceptable probability density for $x \geqslant 0$? Why is $p(x) = 4e^{-2x} - e^{-x}$ not acceptable?

***4** If $p_n = (\frac{1}{2})^n$, show that the probability P that X is a prime number satisfies $6/16 \leqslant P \leqslant 7/16$.

5 If $p(x) = e^{-x}$ for $x \geqslant 0$, find the probability that $X \geqslant 2$ and the approximate probability that $1 \leqslant X \leqslant 1.01$.

6 If $p(x) = C/x^3$ is a probability density for $x \geqslant 1$, find the constant C and the probability that $X \leqslant 2$.

7 If you choose x completely at random between 0 and π, what is the density $p(x)$ and the cumulative density $F(x)$?

†Joe DiMaggio's 56-game hitting streak was much more improbable—I think it is statistically the most exceptional record in major sports.

In 8–13 find the mean value $\mu = \Sigma\, np_n$ or $\mu = \int xp(x)\, dx$.

8 $p_0 = 1/2,\ p_1 = 1/4,\ p_2 = 1/4$

9 $p_1 = 1/7,\ p_2 = 1/7,\ \ldots,\ p_7 = 1/7$

10 $p_n = 1/n!e \quad (p_0 = 1/e,\ p_1 = 1/e,\ p_2 = 1/2e,\ \ldots)$

11 $p(x) = 2/\pi(1+x^2), \quad x \geqslant 0$

12 $p(x) = e^{-x}$ (integrate by parts)

13 $p(x) = ae^{-ax}$ (integrate by parts)

14 Show by substitution that

$$\int_{-\infty}^{\infty} e^{-x^2/2\sigma^2}\, dx = \sqrt{2}\,\sigma \int_{-\infty}^{\infty} e^{-u^2}\, du = \sqrt{2\pi}\,\sigma.$$

15 Find the cumulative probability F (the integral of p) in Problems 11, 12, 13. In terms of F, what is the chance that a random sample lies between a and b?

16 Can-Do Airlines books 100 passengers when their plane only holds 98. If the average number of no-shows is 2, what is the Poisson probability that someone will be bumped?

17 The waiting time for a bus has probability density $(1/10)e^{-x/10}$, with $\mu = 10$ minutes. What is the probability of waiting longer than 10 minutes?

18 You make a 3-minute telephone call. If the waiting time for the next incoming call has $p(x) = e^{-x}$, what is the probability that your phone will be busy?

19 Supernovas are expected about every 100 years. What is the probability that you will be alive for the next one? Use a Poisson model with $\lambda = .01$ and estimate your lifetime. (Supernovas actually occurred in 1054 (Crab Nebula), 1572, 1604, and 1987. But the future distribution doesn't depend on the date of the last one.)

20 (a) A fair coin comes up heads 10 times in a row. Will heads or tails be more likely on the next toss?

(b) The fraction of heads after N tosses is α. The expected fraction after $2N$ tosses is ______.

21 Show that the area between μ and $\mu + \sigma$ under the bell-shaped curve is a fixed number (near 1/3), by substituting $y =$ ______:

$$\int_{\mu}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, dx = \int_0^1 \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy.$$

What is the area between $\mu - \sigma$ and μ? The area outside $(\mu - \sigma, \mu + \sigma)$?

22 For a *yes-no* poll of two voters, explain why

$$p_0 = (1-p)^2,\ p_1 = 2p - 2p^2,\ p_2 = p^2.$$

Find μ and σ^2. N voters give the "*binomial distribution*."

23 Explain the last step in this reorganization of the formula for σ^2:

$$\sigma^2 = \int (x-\mu)^2 p(x)\, dx = \int (x^2 - 2x\mu + \mu^2) p(x)\, dx$$
$$= \int x^2 p(x)\, dx - 2\mu \int xp(x)\, dx + \mu^2 \int p(x)\, dx$$
$$= \int x^2 p(x)\, dx - \mu^2.$$

24 Use $\int (x-\mu)^2 p(x)\, dx$ and also $\int x^2 p(x)\, dx - \mu^2$ to find σ^2 for the ***uniform distribution***: $p(x) = 1$ for $0 \leqslant x \leqslant 1$.

25 Find σ^2 if $p_0 = 1/3,\ p_1 = 1/3,\ p_2 = 1/3$. Use $\Sigma (n-\mu)^2 p_n$ and also $\Sigma\, n^2 p_n - \mu^2$.

26 Use Problem 23 and integration by parts (equation 7.1.10) to find σ^2 for the *exponential distribution* $p(x) = 2e^{-2x}$ for $x \geqslant 0$, which has mean $\frac{1}{2}$.

27 The waiting time to your next car accident has probability density $p(x) = \frac{1}{2}e^{-x/2}$. What is μ? What is the probability of no accident in the next four years?

28 With $p = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$, find the average number μ of coin tosses by writing $p_1 + 2p_2 + 3p_3 + \cdots$ as $(p_1 + p_2 + p_3 + \cdots) + (p_2 + p_3 + p_4 + \cdots) + (p_3 + p_4 + p_5 + \cdots) + \cdots$.

29 In a poll of 900 Americans, 30 are in favor of war. What range can you give with 95% confidence for the percentage of peaceful Americans?

30 Sketch rough graphs of $p(x)$ for the fraction x of heads in 4 tosses of a fair coin, and in 16 tosses. The mean value is $\frac{1}{2}$.

31 A judge tosses a coin 2500 times. How many heads does it take to prove with 95% confidence that the coin is unfair?

32 Long-life bulbs shine an average of 2000 hours with standard deviation 150 hours. You can have 95% confidence that your bulb will fail between ______ and ______ hours.

33 Grades have a normal distribution with mean 70 and standard deviation 10. If 300 students take the test and passing is 55, how many are expected to fail? (Estimate from Figure 8.12b.) What passing grade will fail 1/10 of the class?

34 The average weight of luggage is $\mu = 30$ pounds with deviation $\sigma = 8$ pounds. What is the probability that the luggage for 64 passengers exceeds 2000 pounds? How does the answer change for 256 passengers and 8000 pounds?

35 A thousand people try independently to guess a number between 1 and 1000. This is like a lottery.

(a) What is the chance that the first person fails?

(b) What is the chance P_0 that they all fail?

(c) Explain why P_0 is approximately $1/e$.

36 (a) In Problem 35, what is the chance that the first person is right and all others are wrong?

(b) Show that the probability P_1 of exactly one winner is also close to $1/e$.

(c) Guess the probability P_n of n winners (fishy question).