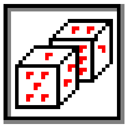
Solutions to Worktext Exercises



Chapter 2

Visualizing Random Processes

Basic Learning Exercises

1. You would expect 5 heads and 5 tails. Samples will vary, but about 90% of the time you will get 2 to 8 heads, inclusive (the same statement applies to tails). There is about a 2% chance of getting at least 9 heads or 1 or fewer heads. Only about 2 times in 1000 will you get 0 or 10 heads. Samples can differ quite a bit.

2. You would expect 25 heads and 25 tails. Samples will vary, but about 90 percent of the time you will get 20 to 30 heads, inclusive (the same statement applies to tails). There is about a 3% chance of 18 or fewer heads or of 32 or more heads (similarly for tails). Chances are less than 1 in 10,000 of seeing 11 or fewer heads or 39 or more heads. You would expect less *relative* variation in the larger sample.

3. The similarities that do exist demonstrate the regularity in the process. Typically, the samples of 10 tosses will appear somewhat different, while the samples of 50 tosses will appear similar. However, any observed differences are due to random variation. It shows that there is regularity in a random process, even though substantial random variation is possible.

4. Since each outcome is equally likely, the result should be 50 zeros (left-side balls) and 50 ones (right-side balls). Considerable variation is possible, but 95% of the time you will get from 42 to 58 zeros, inclusive. Yes, it is like flipping one coin 100 times because the probability of "success" is 0.50.

5. You would expect 30 "successes" (hits) in 100 tries, but the likely range is from 21 to 39. If a batter's average is 0.300 there is no guarantee of getting *exactly* 30 hits in 100 at-bats, due to random variation. This applies to your experiment, as well.

6. The most frequent outcome is 2. It is like flipping four coins because the probability of falling right ("success" or "heads") is the same as the probability of falling left ("failure" or "tails").

7. If you set the probability of falling right to 0.9, you would expect 9 "successes," but 10 is almost as likely. Fewer than 7 would be unusual. This is unlike the coin experiment, because the outcomes are asymmetric.

8. Results will vary, but theoretically 1000 balls should best, because there is more regularity in large samples.

Balls	P(X=0)	P(X=1)	P(X=2)	P(X=3)	P(X=4)	P(X=5)
100	.280	.430	.200	.090	.000	.000
500	.310	.426	.206	.052	.006	.000
1000	.317	.421	.193	.060	.009	.000

9. Relative frequencies make it easier to estimate empirical probabilities, but they suppress your awareness of the actual frequencies and may make the results seem more abstract.

Intermediate Learning Exercises

10. In six throws, you would expect one of each, but some outcomes may not occur at all while others occur 2 or even 3 times. In 12 throws, you would expect two of each outcome, but the probability of some “missing" events is still substantial. Because the sample is small, it is inappropriate to infer that the die is unfair.

11. The distribution should be uniform. For a fair die, we would expect 16 or 17 of each outcome ($1/6 \times 100$), but variation within the range 10 through 24 is not unusual. If we saw 9 or fewer or 25 or more of any outcome, we might doubt on the die’s fairness. 100 throws is a moderately large sample, but a larger sample is always preferred.

12. For a fair die, we would expect 167 of each outcome ($1/6 \times 1000$), but variation within the range 144 through 189 is common. No, variability in frequencies can persist. A large sample makes extreme results less likely, but will not guarantee the expected outcomes.

13. The true distribution is triangular, with expected frequencies of about 3, 6, 9, 12, 15, 18, 15, 12, 9, 6, 3. You generally will get a result that covers the whole domain from 2 to 12. Its theoretical mode is 7, though a sample mode of 6 or 8 would not be surprising. However, its triangular shape may not be apparent. In 1000 throws, the mode of 7 should be unambiguous, and the triangular shape should be more obvious. 100 throws is a rather small sample to feel confident about the shape, but 1000 is much better.

14. The domain is generally from 0 to 5. The mode is either 0 or 1. The histogram is severely right-skewed. Its right tail tapers off rapidly, seeming to hit zero at about 5 defects.

15. Most of the time the number of defects will go from 0 to 5, but you could get 6 (although 7 is shown on the histogram scale it is extremely unlikely). Either 0 or 1 could be modes (both are equally likely). The distribution is skewed to the right (long right tail). Your empirical estimates of the probabilities should be close, given the large sample size.

Sample	$P(X=0)$	$P(X=1)$	$P(X=2)$	$P(X=3)$	$P(X=4)$	$P(X=5)$	$P(X\geq 6)$
1	.355	.389	.183	.052	.017	.001	.003
2	.396	.365	.170	.057	.011	.000	.001
3	.377	.366	.177	.052	.022	.004	.002

16. The number of defects typically goes from 0 to 13, but you may get 14 or even 15. The mode is 4 or 5 (both are equally likely). The histogram is right-skewed, but is closer to being symmetric than before.

17. The number of defects typically goes from 13 to 39, but you may get more extreme values. The histogram no longer seems right-skewed. As the mean increases, the distribution becomes more symmetric and the mode becomes less strong.

Advanced Learning Exercises

18. Because the sample space is {TT, TH, HT, HH} we would expect no heads 25% of the time, one head 50% of the time, and two heads 25% of the time. Yes, with 1000 trials the observed relative frequencies will be fairly close to these percentages (typically within $\pm$ .03).

19. The sample space is too long to list but looks like this: {TTTT, TTTH, TTHT, THTT,HTTT, ..., HHHH}. The expected histogram is symmetric with a domain from 0 to 4 and a strong mode of 2. The probability of P(1) is 0.2500. With a sample of 1000 replications, the relative frequency should be close to 0.25 (typically within $\pm$ .03) but there can be variation.

20. The expected histogram is symmetric with a domain from 0 to 8 and a clear mode of 4. Although the histogram may seem “too spread out” to be called “bell-shaped,” this is because the scale is “stretched” to the screen width. Actually, this symmetric distribution does closely resemble a normal “bell curve.”

21. The domain is 3 through 18. The distribution is symmetric and somewhat bell-shaped. The are mode are 10 and 11, but you could get a mode of 9, 10, 11, or 12. It is possible, though unlikely, to have modes of 8 or 13. The mode is not very clear-cut.

	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Mode(s)1	10, 11	10, 11	10, 11	10, 11	10, 11

22. The domain is from 4 to 24. The distribution is symmetric and somewhat bell-shaped, but is more widely dispersed and has less centrality than the three dice experiment. Logically, the mode ought to be 14 (halfway between 4 and 24). The sample mode is unreliable, even with 1000 repetitions. As the number of dice increases, dispersion increases but the distribution remains symmetric with a peak at the midpoint.

23. This table shows the approximate 98% range that you would expect (based on the Poisson distribution). You should be able to see that the range is increasing, but at a decreasing rate as the mean is increased. The range increases because more possibilities exist in the domain.

Sample	Mean = 4	Mean = 8	Mean = 12	Mean = 16	Mean = 20
Maximum	9	14	20	25	29
Minimum	1	2	5	8	11
Range	8	12	15	17	18