Chapter 4.4 Summarizing a Probability Distribution

Jim Albert and Monika Hu

Chapter 4 Discrete Distributions

Back to Coin-Flipping Example

- Once we have constructed a probability distribution, we can use this to find probabilities.
- What is the chance that Peter will win at least \$5 in this game? Winning "at least \$5" includes the possible values $G = 6$ and $G = 10$.
- One finds the probability of interest by adding the probabilities of the individual values.

$$
\begin{aligned}
P(G \geq 5) &= P(G = 6 \text{ or } G = 10) \\
&= P(G = 6) + P(G = 10) \\
&= \frac{5 + 1}{32} = \frac{6}{32}.
\end{aligned}
$$

Another Probability Computation

- What is the probability Peter wins money in this game? Peter wins money if the gain G is positive and this corresponds to the values $G = 2, 6, 10$.
- So probability that Peter wins money is

$$\begin{aligned} P(\text{Peter wins}) &= P(G > 0) \\ &= P(G = 2) + P(G = 6) + P(G = 10) \\ &= \frac{10 + 5 + 1}{32} = \frac{1}{2}. \end{aligned}$$

Summaries

- It is helpful to compute an "average" of a probability distribution.
- A common measure of "average" is the mean or expected value of X, denoted μ or $E(X)$. The mean is found by

1. Computing the product of a value of X and the corresponding value of the pmf $f(x) = P(X = x)$ for all values of X.
2. Summing the products.

- The formula is:

$$\mu = \sum_x x f(x).$$

Example

The computation of the mean for the Peter-Paul game is illustrated below.

- One sees that the mean of G is $\mu = 0$.

g	$P(G = g)$	$g \times P(G = g)$
-10	$1/32$	$-10/32$
-6	$5/32$	$-30/32$
-2	$10/32$	$-20/32$
2	$10/32$	$20/32$
6	$5/32$	$30/32$
10	$1/32$	$10/32$
SUM	1	0

Interpretation

- How does one interpret a mean value of 0?
- Note that $G = 0$ is not a possible outcome of the game.
- But if Peter and Paul play this game a large number of times, then the value $\mu = 0$ represents (approximately) the mean winnings of Peter in all of these games.

Simulating the Peter-Paul Game (continued)

- The functions `sample()` and `replicate()` were earlier illustrated to simulate this game 1000 times in R.
- Peter's winnings in the different games are stored in the vector `G`. Here is a display of Peter's winnings in the first 100 games:

```
G[1:100]
```

```
##   [1]  -2  10  -6   6  -2   2   2   6   2   6  10  -
##  [19]  -2  -2   6   2  -6   6   2   2   2   2   6
##  [37]  -6   2  -2 -10   6  -6   2  -2   2  -2   2
##  [55]   2  -2   6   2  -6  -2   2  -2   2 -10   2  1
##  [73]   6  -2  -2  -2   6   2  -2  10   2  -6  -2  -
##  [91]  -2  -2  -2  -6   2  -2  -2   2   2   2
```

Compute Sample Mean

- One approximates the mean winning μ by finding the sample mean $\bar{G}$ of the winning values in the 1000 games.

```
mean(G)
```

```
## [1] -0.008
```

- This value is approximately equal to the mean of G, $\mu = 0$.
- If Peter was able to play this game for a much larger number of games, his average winning would be very close to $\mu = 0$.