The Bernoulli Probability Distribution

A random variable with only two outcomes (1-0, true-false, right-wrong, on-off, etc.) is called a **Bernoulli random variable**, denoted as $X\tilde{}$Bernoulli(p). For computing purposes we limit the numerical values to 1 for a success and 0 for a failure.

Example 1 *What is the expected value and standard deviation for a Bernoulli random variable with probability of success p?*

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

$$\begin{aligned} E(X) &= \sum_{x \in X} x * p(x) \\ &= 1 * p + 0 * (1-p) \\ &= p \end{aligned}$$

$$\begin{aligned} \sigma^2 &= \sum_{x \in X} (x - E(X))^2 * p(x) \\ &= (1-p)^2 * p + (0-p)^2 * (1-p) \\ &= (1-p)(1-p)p + p^2(1-p) \\ &= (1-p)p[(1-p)+p] \\ &= p(1-p) \\ \textit{Thus, } \sigma &= \sqrt{p(1-p)} \end{aligned}$$

Example 2 *It estimated 35% of students in the MSAS program work during the day. What is the standard deviation for this Bernoulli random variable?*
$\sigma = \sqrt{p(1-p)} = \sqrt{.35(1-.35)} = 0.476\,97$.

Example 3 *What percentages of MSAS students with day jobs fall within two standard deviations of the mean?*
$\mu \pm 2\sigma = .35 \pm 2 * 0.476\,97 = (-0.603\,94, 1.303\,9)$

Example 4 *A fair coin is flipped 100 times. What percentage of heads falls within one standard deviation of the mean?*
This is $X\tilde{}$Bernoulli(0.5). So $\mu = p = 0.5$ and $\sigma = \sqrt{p(1-p)} = \sqrt{.5(1-.5)} = 0.5$.

Remark 5 *Note that these intervals cover a lot of ground. That must be the case since the only outcomes for a single observation are 0 and 1.*

0.1 Exercises

Navidi 4.1: 1, 3, 5

The Binomial Probability Distribution

Definition 6 *An experiment consisting of repeated trials has a* ***Binomial Probability Distribution*** *if and only if*

1. *there are only two outcomes for each trial (success or failure) and the probability p of success is fixed;*
2. *there are a fixed number of trials n;*
3. *the trials are independent.*

Remark 7 *Such a random variable is denoted $X\tilde{}Bin(n, p)$. We use it to count a total number of successes.*

Exercise 8 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. Does this experiment have a binomial probability distribution?*

Exercise 9 *A multiple choice test contains 20 questions. Each question has four or five choices for the correct answer. Only one of the choices is correct. With random guessing, does this test have a binomial probability distribution?*

Exercise 10 *A multiple choice test contains 20 questions. Each question has five choices for the correct answer. Only one of the choices is correct. When a student trying their best, does this test have a binomial probability distribution?*

Exercise 11 *A multiple choice test contains 20 questions. Each question has five choices for the correct answer. Only one of the choices is correct. With random guessing, does this test have a binomial probability distribution?*

Exercise 12 *It is known that 30% of all students at the University of Knowhere live off campus. An experiment consists of randomly selecting students until a commuter student is found. Does this experiment have a binomial probability distribution?*

Exercise 13 *An experiment consists of asking 20 students what size water they prefer: small, medium or large. All choices are equally likely. Does this experiment have a binomial probability distribution?*

Exercise 14 *An experiment consists of testing the functionality of five different flash drives from a shipment of 50,000 of which 100 are known to be defective. Does this experiment have a binomial probability distribution? Technically, no since the selections are made without replacement. However the sample size is so small compared to the size of the population that for all practical purposes we can consider the selections to be independent.*

Remark 15 *If the sample size is less than 5% of the population size then it is reasonable to treat selections made without replacement as independent selection.*

What does knowing that an experiment has a binomial probability distribution buy us? Quite a lot actually.

Theorem 16 *If an experiment X has a binomial probability distribution with n trials and probability of success p then*

1. *the probability of exactly k successes is $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$;*
2. *the mean or expected value is $E(X) = \mu = np$;*
3. *and the variance $V(X) = np(1-p)$ and standard deviation $\sigma = \sqrt{np(1-p)}$.*

Proof. Let X have a binomial probability distribution with n trials and probability of success p.

1. Which of the n trials yield a success? We can pick any k of them in an unordered fashion in $\binom{n}{k}$ ways. We must succeed on those k selections which has a likelihood of p^k. Next we must fail on the remaining $n-k$ trials. Doing so has likelihood $(1-p)^{n-k}$. Thus, the probability of exactly k successes is $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$.
2. Best left for STAT 7020.
3. Best left for STAT 7020.

■

Remark 17 *Our book uses tables of values found in the appendix to compute binomial probabilities. Feel free to use these or the technology/software of your choice. For example, on the TI-83/84 series of calculators, **binompdf**(n, p, k) computes the probability of exactly k successes. The command **binomcdf**(n, p, k) computes the probability of 0 or 1 or 2 or...or k successes.*

Example 18 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. Find the mean and standard deviation for this binomial probability distribution.*

Since $n = 10$ and $p = \frac{1}{2}$, $\mu = 10 * \frac{1}{2} = 5$ and $\sigma = \sqrt{10 * \frac{1}{2} * (1 - \frac{1}{2})} = 1.5811$.

Example 19 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. What is the probability of observing exactly five tails?*

Since $n = 10$, $p = \frac{1}{2}$ and $k = 5$, the probability of observing exactly 5 tails is $p(k = 5) = \binom{10}{5} * (\frac{1}{2})^5 \left(1 - \frac{1}{2}\right)^{(10-5)} = 0.246\,09$

Example 20 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. What is the probability of observing at most three tails?*

Here $n = 10$ and $p = \frac{1}{2}$ but $k = 0, 1, 2$, or 3. These are disjoint cases so we use the addition rule four times.

$p(k \leq 3) = \binom{10}{0} * (\frac{1}{2})^0 \left(1 - \frac{1}{2}\right)^{(10-0)} + \binom{10}{1} * (\frac{1}{2})^1 \left(1 - \frac{1}{2}\right)^{(10-1)} + \binom{10}{2} * (\frac{1}{2})^2 \left(1 - \frac{1}{2}\right)^{(10-2)} + \binom{10}{3} * (\frac{1}{2})^3 \left(1 - \frac{1}{2}\right)^{(10-3)} = 0.171\,88$

Example 21 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. What is the probability of observing at least nine tails?*

Here $n = 10$ and $p = \frac{1}{2}$ but $k = 9$ or 10.

$p(k \geq 9) = \binom{10}{9} * (\frac{1}{2})^9 \left(1 - \frac{1}{2}\right)^{(10-9)} + \binom{10}{10} * (\frac{1}{2})^{10} \left(1 - \frac{1}{2}\right)^{(10-10)} = 1.074\,2 \times 10^{-2}$.

Example 22 *An experiment consists of flipping a fair coin 10 times and counting the number of tails. What is the probability the number of tails falls within two standard deviations of the mean?*

Since $5 \pm 2 * 1.58811$ yields the interval $1.823\,8$ to $8.176\,2$, the probability is $\binom{10}{2} * (\frac{1}{2})^2 \left(1 - \frac{1}{2}\right)^{(10-2)} + \binom{10}{3} * (\frac{1}{2})^3 \left(1 - \frac{1}{2}\right)^{(10-3)} + \binom{10}{4} * (\frac{1}{2})^4 \left(1 - \frac{1}{2}\right)^{(10-4)}$
$+ \binom{10}{5} * (\frac{1}{2})^5 \left(1 - \frac{1}{2}\right)^{(10-5)} + \binom{10}{6} * (\frac{1}{2})^6 \left(1 - \frac{1}{2}\right)^{(10-6)} + \binom{10}{7} * (\frac{1}{2})^7 \left(1 - \frac{1}{2}\right)^{(10-7)}$
$+ \binom{10}{8} * (\frac{1}{2})^8 \left(1 - \frac{1}{2}\right)^{(10-8)} = 0.978\,52$.

Problem 23 *Pandora radio offers a thumb print station which consists solely of songs given a thumbs up by the account user. This gets tricky when married couples share an account on communal devices (Roku, fire stick, etc.). Of all songs given a thumbs up, Joe DeMaio is responsible for 60% while Sylvia DeMaio is responsible for the remaining 40% (no overlap between song choices). Of the next 20 songs played, what is the probability that at least half emanate from Sylvia's songs?*

Problem 24 *In 2017, 71% of all full-time KSU undergraduates received some type of need-based financial aid (https://www.usnews.com/best-colleges/kennesaw-state-university-1577). Twenty students are selected at random.*

1. *Find the mean and standard deviation for the number of students who received need-based financial aid.*

2. *Compute the probability that exactly fourteen of the twenty students received need-based financial aid.*

3. *Compute the probability that at least fifteen of the twenty students received need-based financial aid.*

4. *Suppose only two of the twenty KSU students on the intramural water-polo team received need-based financial aid. What does that suggest about water-polo.*

What if we do not know the fixed probability of success? Can we estimate it from a sample? Yes.

Example 25 *Dave and Kathy also share a Pandora account but have never kept track of their thumbs up declarations. Of the next 50 songs, only 10 were given a thumbs up by Kathy with the remainder from Dave's list.*
it makes sense to estimate p from the sample data. We will denote the sample estimate as $\hat{p}$ and infer that $p = \hat{p}$. In this case $\hat{p} = .2$ from Kathy's perspective and $\hat{p} = .8$ from Dave's perspective.

How much uncertainty do we have in our estimates of $\hat{p}$? After all, it is just an estimate. For proportions, the uncertainty is $\sigma_{\hat{p}} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

Example 26 *What is the uncertainty in our estimates for Dave and Kathy's thumb print radio preferences? From Kathy's perspective this is $\sigma_{\hat{p}} = \sqrt{\dfrac{.2*(1-.2)}{50}} =$ 5.6569×10^{-2}. Due to symmetry about p in our formula for $\sigma_{\hat{p}} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$, from Dave's perspective we get the very same $\sigma_{\hat{p}} = \sqrt{\dfrac{.8*(1-.8)}{50}} = 5.6569 \times 10^{-2}$.*

Example 27 *Within two levels of uncertainty (standard errors), what percentage of thumbs-up songs emanate from Dave?*
$.8 \pm 2 * .057 = (.686, 0.914)$

Example 28 *What if we think our level of uncertainty is too large. Say we want to shrink the* .057 *down to under* 0.03*. We can't change the estimate of proportion. However, we do have control over the sample size. So, solve for n* *in* $\sigma_{\hat{p}} = \sqrt{\dfrac{.8*(1-.8)}{n}} \leq .3.$

$$\begin{aligned}
\sqrt{\frac{.8*(1-.8)}{n}} &\leq .03 \\
\frac{.8*.2}{n} &\leq .0.0009 \\
.16 &\leq .0009n \\
\frac{.16}{.0009} &\leq n \\
177.78 &\leq n
\end{aligned}$$

So, to decrease the uncertainty rate to under 0.03, increase the sample size to 178. Note that we are not rounding to 178 but rather applying the ceiling function.

0.2 Exercises

Navidi 4.2: 1-4, 7-10, 12, 15-18