|  |  |  |
| --- | --- | --- |
| * **B = Binary Outcomes**: Each individual trial may be considered either **success** or **failure**. (In particular, it is not returning number other than 0 or 1. * **I = Independent**: Each individual trial does not affect the outcome of another. * **N = Fixed Sample Size**: The number of trials is pre-specified, and does not depend on the outcome of those trials. * **S = Same Probability**: Each individual trial has the same probability of success as any other     Binomial coefficient chose(n,k) (n over k) chose(3, 1) ##3 | D~Binom(3,1/6) P(D=1)  roll1 = c(F, T, F, F, T, T, F, T),  roll2 = c(F, F, T, F, T, F, T, T),  roll3 = c(F, F, F, T, F, T, T, T)  p = 1/6  k = 1  n = 3  round(p^k \* (1-p)^(n-k), 3)  ## [1] 0.116  # They match!  three\_rolls %>% filter(D == 1) | ## # A tibble: 8 × 6  ## roll1 roll2 roll3 prob\_str prob D  ## <lgl> <lgl> <lgl> <chr> <dbl> <int>  ## 1 FALSE FALSE FALSE 5/6 \* 5/6 \* 5/6 0.579 0  ## 2 TRUE FALSE FALSE 1/6 \* 5/6 \* 5/6 0.116 1  ## 3 FALSE TRUE FALSE 5/6 \* 1/6 \* 5/6 0.116 1  ## 4 FALSE FALSE TRUE 5/6 \* 5/6 \* 1/6 0.116 1  ## 5 TRUE TRUE FALSE 1/6 \* 1/6 \* 5/6 0.023 2  ## 6 TRUE FALSE TRUE 1/6 \* 5/6 \* 1/6 0.023 2  ## 7 FALSE TRUE TRUE 5/6 \* 1/6 \* 1/6 0.023 2  ## 8 TRUE TRUE TRUE 1/6 \* 1/6 \* 1/6 0.005 3 |
| **dbinom(k, size = n, prob = p),** **P(Binom(n,p)=k),****dbinom(k,n,p)**  ***mean***Binom(n,p)*is* ***n\*p****.*  *standard deviation* |  | |
| Simulation random gen **rbinom(n, size, prob) n-values** | rbinom(3, size = 10, prob = 0.5) ## [1] 5 3 5 | |
| random\_data = rbinom(1000000, size = 100, prob = 0.5)  mean(random\_data)  ## [1] 49.99412  var(random\_data)  ## [1] 25.00711 | dbinom(51, 100, 0.5)  ## [1] 0.07802866  mean(random\_data == 51)  ## [1] 0.078038 | |
| **gbinom(n,p) graphs binom(n,p)**  gbinom(n = 2, p = 0.4) +  geom\_vline(xintercept = 2\*0.4, linetype = "dashed", color = "red")  ggplot() + geom\_binom\_density() |  | |
| Cumulative probability CDF  d<dist>, p<dist>, q<dist>,r<dist> | *pbinom(q, size, prob)* P(Binom(size,prob)≤q)  # What is the probability of getting four or fewer heads from 10 coin flips?  **pbinom(4, n, p)**  ## [1] 0.3769531 | |
| # The manual alternative to **pbinom(4, n, p)**:  dbinom(0, n, p) + dbinom(1, n, p) + dbinom(2, n, p) + dbinom(3, n, p) + dbinom(4, n, p)  sum(dbinom(0:4, n, p)) | Probability to the right **1 - pbinom(x, n, p)**  # This is P(X > 4), or P(X >= 5)  1 - pbinom(4, n, p)  # This is P(X > 4), or P(X >= 5)  **pbinom(4, n, p, lower.tail = FALSE) P(X>4)** | |
| P(a <= Binom(n, p) <= b) P(c<X<d) P(c+1<=X<=d−1)  pbinom(b, n, p) - pbinom(a - 1, n, p). | **qbinom(p, size, prob)**  tells you how large q needs to be for P(Binom(size,prob)≤q) to be at least p. It is the inverse of pbinom(q, size, prob).  # To capture at least 70% of the probability in Binom(10, 0.5), how large do we need to make x?  **qbinom(0.7, n, p) quantile q for probability 0.7** | |

|  |  |
| --- | --- |
| X∼N(μ,σ)  **dnorm(x, mean, sd)**  gives you the height of the density function of N(mean,sd) at x*.* |  |
| **qnorm(p,mean,sd)** How large do we need to make q so P(N(100,25)<q) | qnorm(0.7, 100, 25)  ## [1] 113.11 |
| We can tell by considering **how many standard deviations away from their mean** each point was!  Z=(x−μ)/σ | **If X∼N(μ,σ)**  **then (X−μ)/σ=Z∼N(0,1)** |
| X∼N(100,25) Find P(X<80) **pnorm(z)**  mu = 100  sigma = 25  x = 80  z = (x-mu)/25  ## Using X directly  **pnorm(x, mu, sigma)** | Central Limit theorem  Consider sampling n points from any distribution D, which has expected value μ  and standard deviation σ, and then taking the mean of those n  points as a random variable X¯  **Then, X¯∼N(μ,σ/√n)** |
|  |  |
| Point estimate p^    P^ standard error of distribution | **Confidence interval**   * Let’s say we want to be **95% confident** that our region contains the true value of p. * The width of our confidence interval should then be **the width of the region which contains 95% of the area on the sampling distribution**. |
|  | **qnorm(0.975, true\_p, stdError)**  ## [1] 0.615596  The general adjustment here is that if your confidence level is α, then the right endpoint is the α+(1−α)/2 quantile of the sampling distribution.  **qnorm(conf.level + (1 - conf.level) / 2, true\_p, stdError)**  **## [1] 0.615596**  **true\_p + qnorm(conf.level + (1 - conf.level) / 2) \* stdError**  **## [1] 0.615596** |
|  | **Increase n** |
|  |  |
| **The Wald Adjustmen SE(p^)** |  |
| # Three inputs; successes, sample size, confidence level  x = 60  n = 90  conf.level = 0.95  # Calculate Wald confidence interval  phat = x/n  se = sqrt(phat \* (1 - phat)/n) # Wald adjustment replaces p with phat  moe = qnorm(conf.level + (1 - conf.level)/2) \* se    left = phat - moe  right = phat + moe    c(left, right)  ## [1] 0.5692753 0.7640581 | * We could say that we are **95% confident** that the **true value of**p is between 56.9% and 76.4%, with the Wald adjustment. * **We are 95% confident that the true probability of a chimpanzee making a prosocial choice with a partner is between 56.9% and 76.4% (with the Wald adjustment).**   **The \*\*Agresti-Coull\* adjustment adds two successes and two failures to the dataset artificially. This changes X to X+2 and n to n+4**  **These changes carry through to the point estimate, which is passed into the standard error**. |
| # Three inputs; successes, sample size, confidence level  x = 60  n = 90  conf.level = 0.95  # Calculate AC confidence interval  x\_adjusted = x+2 # Added two successes  n\_adjusted = n+4 # Added four data points total  phat = x\_adjusted/n\_adjusted  se = sqrt(phat \* (1 - phat)/n\_adjusted) # Changes carry through to p\_hat and n\_adjusted  moe = qnorm(conf.level + (1 - conf.level)/2) \* se    left = phat - moe  right = phat + moe    c(left, right)  ## [1] 0.5637829 0.7553660 | **We are 95% confident that the true probability of a chimpanzee making a prosocial choice with a partner is between 56.4% and 75.5% (with the Agresti-Coull adjustment).** |
| **Step 1: Model Statement****Defining Variables** Let X be the number of observed prosocial choices among **n=90** chimpanzee A trials with a partner  **Model statement**  X∼Binom(90,p) **Model Assumptions**  * We may declare that X follows a binomial distribution only if we accept the four binomial assumptions: * **B**: Chimpanzee A either makes the prosocial choice or does not. * **I**: Each trial can be assumed to be independent of each other. * **N**: We have a fixed number of trials, **n=610**. * **S**: Each trial can be assumed to have the same true p of “success”. | Step 2: State Hypotheses The null hypothesis is always of the form **true parameter value = some number**, and captures the idea that there is no meaningful pattern, or no meaningful relationship, et cetera.  The alternative hypothesis is of the **form true parameter value \neq that same** **number or true parameter value > that same number or true parameter value <** **that same number**. It captures the idea that there is a meaningful pattern or relationship in your data.   * With the chimpanzee data, we can test the null hypothesis that p=0.5; the idea that the chimps are picking randomly between the colors, they really aren’t considering their partner at all. * The alternative hypothesis, or the idea that the chimps ARE picking prosocially more often than random, is p>0.5. * We write in statistical notation:   **H0:p=0.5**  **HA:p>0.5** |
| Step 3: Test Statistic and Null Distribution To find statistical evidence AGAINST the null hypothesis:  First, **Assume the null hypothesis is true**. In this case, assume p=0.5  Second, we need a test statistic; some statistic from the sample for which we know the probability distribution. Then, we plug the null value of the parameter into it, in this case replacing p with 0.5  When we plug in the null value, we call that resulting distribution the null distribution.  The test statistic comes straight from our model: we know that the observed number of successes X∼Binom(90,p)  When we assume p=0.5, this turns into X∼Binom(90,**0.5**)  That is our null distribution. | Step 4: Identify Relevant Outcomes from Data + Alt. Hyp.  We now bring in our observed data; in this case **X=60**  We need to consider both the observed data and the alternative hypothesis (remember, the statement we actually want to prove) to determine the following set:  Which values of the test statistic that are as or less likely than the one we observed (on the null distribution from step 3) also constitute evidence for the alternative hypothesis?  Let’s look at the null distribution below, and annotate the observed value of our test statistic: null\_p = 0.5  **gbinom(90, null\_p, scale = TRUE) +**  **geom\_vline(xintercept = 60, color = "red")** |
|  | - The values which are **as or less likely** than our observed 60 are those greater than or equal to 60, AND those less than or equal to 30. (Because the null distribution among is symmetric, p=0.5  - Recall that the alternative hypothesis was p>0.5. The lower outcomes don’t constitute evidence for a large p, but the higher outcomes do. Therefore, our set of interest is **values greater than or equal to 60**. |
| **Step 5: Calculate P-Value**  The p-value of a hypothesis test is the probability of observing a value of the test statistic as or more extreme than the one we did, on the null distribution, with “extreme” defined by the null hypothesis.  -Note: p-value is the term specifically for the probability calculated in a hypothesis test. This is sometimes confused with “the value of p”, the true parameter value we are interested in.  This is the probability of observing an outcome in the set identified in step 4. Note that since this is a probability, it should be between 0 and 1. Some p-values will get very low, but should never be negative or greater than 1.  On Binom(90,0.5)  (the null distribution), the probability of getting 60 or greater should be small, according to the picture in step 4.  # X ~ Binom(90, 0.5):  # P(X >= 60) = P(X > 59) = 1 - P(X <= 59) = 1 - pbinom(59, 90, 0.5)  1 - **pbinom(59, 90, 0.5)**  **## [1] 0.001030133** | **Step 6: Interpret In Context**  We have never proven the null hypothesis false, nor the alternative hypothesis true. We have merely succeeded or failed in finding strong evidence against the null.  we have found strong evidence a “low enough” p-value? common standard, for  p-values < 0.05 constitute strong evidence AGAINST the null. “statistically significant”.  **There is strong evidence that chimpanzee A would choose the prosocial choice more than half the time in the long run.**  This is not a statement about **truth,** this is a measure of **the statistical evidence against the null hypothesis.**  **If we had instead observed a p-value above 0.05, we would say that we have failed to find evidence against the null hypothesis.**  **In context, we might say that We fail to find evidence that chimpanzee A is not picking randomly, or that The data is consistent with the idea that chimpanzee A is picking randomly. (p = 0.001, single proportion binomial test)** |
| Hypothesis Testing: TL;DR  * Step 1: State the statistical **model** for the data. * Step 2: State your **hypotheses**. H0=0 should be a statement about a parameter with =, and HA should be another statement about that same parameter with >, <, or ≠. * Step 3: Identify your **test statistic**, and its **null distribution** (by plugging in the H0=0 statement). * Step 4: Identify the set of relevant **outcomes**; those which are as or less likely than your observed value of the test statistic on the null distribution, AND constitute evidence for HA. * Step 5: Calculate the **p-value**, the total probability of the outcomes from step 4 on the null distribution. * Step 6: Interpret the p-value in context using the 0.05 threshold * random card from a deck, determine what suit the one of four suits, each with equal probability; “hearts”, “diamonds”, “spades”, or “clubs”). * Repeated 200 times, got the right suit 57 times.   *Conduct a hypothesis test to assess the evidence for the psychic’s ability.* | **Step 1: Model and Assumptions**  Let X **n=200** trials.  Let p be the true probability of getting the suit right on an individual trial.  X∼Binom(200,p)  BINS assumptions; **B**: each trial is right or wrong, **I**: the trials are independent, **N**: There is a fixed sample size, n=200, and **S**: each trial has the same probability of success, p.  **Step 2: State Hypotheses**   * The null hypothesis captures the idea that picking randomly.   H0:p=0.25   * The alternative hypothesis captures the thing that you seek to prove; that the psychic truly has a better-than-random chance of getting it right.   HA:p>0.25  **Step 3: Test Statistic and Null Distribution**  • We know that X∼Binom(200,p). Under the null hypothesis which states p=0.255, X∼Binom(200,p)  • The observed data is X = 57.  Step 4: Identify Relevant Outcomes  gbinom(200, 0.25, scale = TRUE) +  geom\_vline(xintercept = 57, color = "red") |
| **This null distribution is NOT symmetric (p=0.25, Binomial distribution is only symmetric when p=0.5).**  **- identify all outcomes as or less likely than X = 57**  **dbinom(57, 200, 0.25)**  **## [1] 0.0330775** | **Visually, we expect the threshold to be around X = 43, but we have to confirm:**  **tibble( x = 38:45, prob = dbinom(x, 200, 0.25))**  **## # A tibble: 8 × 2**  **## x prob**  **## <int> <dbl>**  **## 1 38 0.00934**  **## 2 39 0.0129**  **## 3 40 0.0173**  **## 4 41 0.0226**  **## 5 42 0.0285**  **## 6 43 0.0349**  **## 7 44 0.0415**  **## 8 45 0.0479** |
| X = 43 is more likely than X = 57, so it is NOT included in this set of outcomes. X = 42 and below are the outcomes of interest.Thus, the set of all outcomes as or less likely than X=57on Binom(200,0.25) **is X<=42 and X>=57**However, only X>=57 constitutes evidence for alternative hypothesis, HA:p>0.25 | **Step 5: Calculate P-Value**We wish to calculate **P(Binom(200,0.25)>=57)**We will do this by calculating **1−P(Binom(200,0.25)<=56)** # lower.tail = FALSE is same as 1 - pbinom(...)  pbinom(57 - 1, 200, 0.25, lower.tail = FALSE)  ## [1] 0.1445431  Step 6: Interpret in Context  We fail to find strong evidence that picking is not randomly. (p = 0.14, single proportion binomial test) |
|  |  |
|  |  |
|  |  |

|  |  |
| --- | --- |
| How many standard deviations is the 0.72 quantile of a N(μ,σ) distribution above the mean? In other words, if x is the 0.72 quantile, for what value z does x=μ+zσ  F(x)=P(X≤x) | N(0,1) quantil 0.72z= F(z)=0.72 z= F-1(0.72)= 0.553, 0.72= μ+0.553σ  σ =(0.72- μ)/0.553  0.5658. |
| 2.Heights in a population of American adult males are approximately normal with a mean of 69.7 inches and a standard deviation of 3.1 inches.  -2a. What proportion of American adult males are taller than two meters tall? (One meter equals 39.37 inches.)  -2b. What is the 95th percentile of American adult male height?  -2c. If we took a random sample of 250 men, how likely is it that the sample mean of their heights would be between 69.5 and 70.5 inches? | 1m=39.37 inches 2m=78.74 inches  2a. **pnorm(**78.74**,** 69.7**,** 3.1**)**  2.b **qnorm(0.95,** 69.7**,** 3.1**)**  2.c (200, 69.7**,** 3.1)  μ= 69.7 σ=3.1/√200  pnorm(70.5,69.7,3.1/√200) - pnorm(70.5,69.5,3.1/√200) |
| 3.Suppose you are playing a coin flipping game with a friend, where you suspect the coin your friend provided is not a fair coin. In fact, you think the probability the coin lands heads is less than 0.5. To test this, you flip the coin 100 times and observe the coin lands heads 35 times.  -3a. If you assume the coin is fair (i.e., the probability of the coin landing heads is 0.5), what is the probability of observing 35 heads or fewer, calculated using an exact model?  -3b. Calculate the previous probability, but use a normal approximation to achieve a numerical value. What is the relative error in this approximation? (Relative error is the absolute difference divided by the exact value, times 100%.)  -3c. How small would p need to be (rounded to the nearest 0.01) for the probability of observing 35 or fewer heads to be at least 0.05?   * Does it seem plausible that the coin is fair? Briefly explain. |  |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Normal distribution** X∼Normal (μ,σ2)  Mediana=mean, points of inflection, steepest tangent | | |  |
| μ−σ , μ+ σ 68%, μ−2σ , μ+2 σ 95%, μ−3σ , μ+3 σ 99.7%  μ−σ (μ+σ) is about 16%. μ−2σ (μ+2σ) is about 2.5%. | | | F(x)=P(X≤x) discrete P(X<3) and P(X≤3) differ by P(X=3) |
| * left tail: P(Z<z)=Φ(z) * right tail: P(Z>z)=1−Φ(z) * finite interval: P(a<Z<b)=Φ(b)−Φ(a) * outer area: P(|Z|>a)=2Φ(−a) is a>0 and 1s a<0. | | |  |
| * Central limit theorem * The random variables X1,…,Xn are independent and drawn from a distribution F where E(Xi)=μ and Var(Xi)=σ2. * The first two statements are true for any distribution F (where the mean and variance exist). * The accuracy of the approximation to the normal distribution depends both on the size of n and how close F itself is to normal.   + If F is normal, then the distribution of ¯X¯ is also normal for any n. * The strength of skewness in F is the single factor which most strongly affects the quality of the approximation.   + A distribution F with very strong skewness will typically require a larger n (maybe much larger n) for the approximation to be accurate. | | | 1. The sample mean is a random variable defined as 2. If n is large enough, then distribution of ¯X is approximately normal |
| Cumulative funct pnorm()  N(50,10) *## directly*  **pnorm(40, 50, 10)**  [1] 0.1586553 | **standardization**  **pnorm((40 - 50)/10)**  [1] 0.1586553 | **density function**  **dnorm(40, 50, 10)**  [1] 0.02419707 |  | |
| The quantile function, qnorm(), **qnorm(0.9, 50, 10) [1] 62.81552** | | | Random samples, rnorm().rnorm(15, 50,10)  [1] 41.27094 42.13957 76.66895 46.20586 52.95887 43.18090 38.53184 41.31807 [9] 39.57165 29.03125 52.20425 35.74484 57.34440 44.16928 48.15942 |
| **Graphing Normal Distributions compatible with ggprob**   * gnorm() for creating a graph of a normal density * geom\_norm\_density() to add a normal density to a plot * geom\_norm\_fill() to shade in part of a normal density | | | gnorm(50, 10)  gnorm(50, 10, a = 70, color = "hotpink") |
| ggplot() +  geom\_norm\_density(100,10, color = "blue") +  geom\_norm\_density(120,10, color = "red") +  geom\_norm\_density(110,20, color = "violet") +  geom\_hline(yintercept = 0) +  ylab("density") +  theme\_bw() | | |  |
| ggplot() +  geom\_norm\_fill(100,10, fill = "cyan", alpha = 0.5) +  geom\_norm\_fill(120,10, fill = "magenta", alpha = 0.5) +  geom\_hline(yintercept = 0) +  ylab("density") +  theme\_bw() | | |  |
| gnorm(500,100, color = "black") +  geom\_norm\_fill(500,100, a = 650, fill = "firebrick") + *## right tail*  geom\_norm\_fill(500,100, b = 350, fill = "firebrick") + *## left tail*  geom\_norm\_fill(500,100, a = 350, b = 650, fill = "gray") + *## center*  geom\_hline(yintercept = 0) +  ylab("density") +  theme\_bw() | | |  |
|  | | |  |
|  | | |  |