Econ 270 Lecture 6

Sam Gifford

2025-02-17

Random Variables

- We've already learned how to calculate probability and defined events, but sometimes it makes sense to perform operations on an entire distribution
- A **Random Variable** contains all of the information about a probability distribution
  - Mathematically it just maps events to it's sample space
  - For this class, we care more about the operations we can perform on it than it's definition

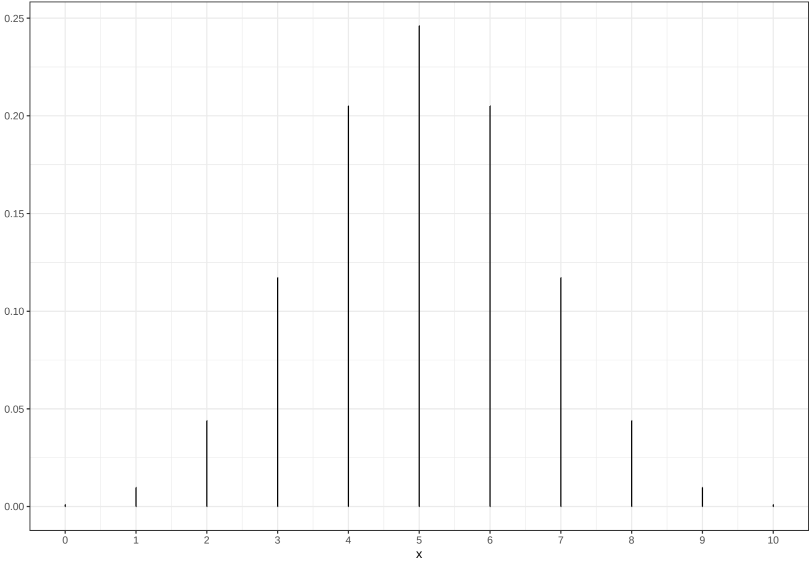
Random Variables

- Toss a coin 10 times and let the random variable X denote the number of heads that occur
- The shorthand $X = 5$ denotes the event that 5 heads occurs. $P(X = 5) = \binom{10}{5}(\frac{1}{2})^{10}$
- We can similarly denote $P(X = 0), P(X = 10), P(X \geq 3)$, etc
- Generically, we use $P(X = x)$ to denote the probability of obtaining x coin flips
  - $P(X = x) = \binom{10}{x}(\frac{1}{2})^{10}$

Random Variables: probability mass function

- For a discrete random variable we use the shorthand $p(x) = P(X = x)$ to denote the probability that an event occurs
- When graphed, the pmf is anologous to the empirical histogram
- What is $p(x)$ for a random variable that represents the number of heads in 10 coin flips?

pmf: 10 coin flips



Binomial Distribution

- We can fully specify a distribution by giving it's probability mass function.
- The binomial distribution is the distribution of getting x successes from n independent events each with probability p
- Coin flipping is a specific case when p=0.5:
- $p(x) = \binom{n}{x} p^x (1-p)^{(n-x)}$
  - If you're good at algebra, you can confirm $\sum_{x=0}^{x=n} p(x) = 1$

Expectation

- We can also calculate an average of a random variable. We call this the expectation and use the operator E
  - $\mu = E[X]$
- The mean is an average weighted by the probability:
  - $E[X] = \sum_x xp(x)$
- We can similarly calculate $var(X) = E[(X - E[X])^2)$
  - $= \sum_x (x - \mu)^2 p(x)$

Expectation Example

- Let X be a binomial variable with n=3, p=.5 ($X \sim Binom(3, 0.5)$)
  - This is the number of heads in 3 tosses of a coin
- Calculate $E[X]$

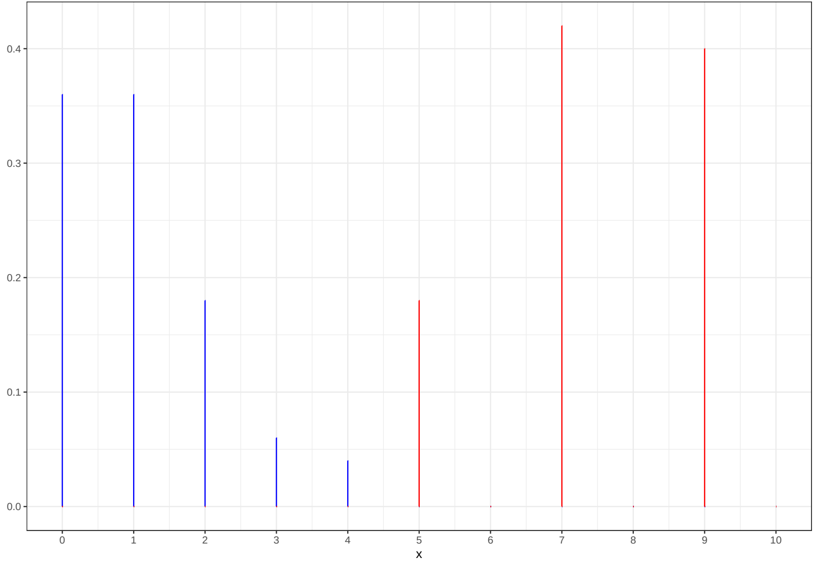
Random Variable Operations

- We partly define random variables to introduce the concepts of pmf and expectation, but mostly we want to do operations on random variables
- Let X_1 be the toss of a fair coin, and X_2 be the toss of a second fair coin. What is $X_1 + X_2$?
  - $Y = X_1 + X_2$ is its own random variable with a pmf and expectation
  - Expectation is easy to calculate. Variance is slightly harder. The pmf is much harder.

Expectation Example

- Let X be a censored Poisson random variable defined by the following pmf:
- $p(0) = .36, p(1) = .36, p(2) = .18, p(3) = .06, p(4) = .04$
- Calculate E[X]
- Let $Y = (X-2)^2 + 5$. Calculate $E[Y]$

Expectation Example: pmf



The Bernoulli Random Variable

- First, let's define X_1 as the success of an event with probability p
  - $X = 1$ if there is a success, $X = 0$ otherwise
- What is $p(x)$?
- What is $E[X]$?

Random Variable Addition: Expectation

- First, note that X_1 and X_2 are identical, so e.g. $E[X_1] = E[X_2]$
- To get $E[X_1 + X_2]$, use linearity. General formula: $E[\sum \alpha_i X_i] = \sum \alpha_i E[X_i]$
  - $E[X_1 + X_2] = E[X_1] + E[X_2]$

Random Variable Addition: Example

- Let X be a binomial variable with n=10, p=.5 ($X \sim Binom(10, 0.5)$)
  - This is the number of heads in 10 tosses of a coin
- Calculate $E[X]$

Random Variable Addition: Example

- In Dungeons and Dragons, the damage of the spell meteor swarm is calculated by rolling 20 6-sided dice and adding up all of the values.
- Calculate the average damage dealt by meteor swarm

Random variable Variance

- For variance, we have $E[(X - E[X])^2]$. First we can distribution the inner part, then we apply linearity of expectation:
  - $var(X) = E[X^2] - E[X]^2$
  - This means that $var(X + Y) \neq var(X) + var(Y)$
- General formula: $var(aX + bY) = a^2 var(X) + b^2 var(Y) + 2cov(X, Y)$
  - If the variables are independent, the last term is 0

Bernoulli Variance

- Let X be a Bernoulli random variable with probability p
- Calculate var(X)

Binomial Variance

- Let X be a binomial variable with n=3, p=.5 ($X \sim Binom(3, 0.5)$)
- Calculate $var(x)$

Variance example

- What is the variance of X, where X represents the roll of a fair 6-sided die?
- What is the standard deviation of meteor swarm?

Variance example

- Suppose that your total grade is the sum of your quiz scores and an exam score
- Aggregate quiz scores have variance of 340 while exam scores have variance of 3100
- The correlation between quiz and exam scores is .4
- Calculate the variance of total grade

Variance Example

- Your quiz subtotal is made up of 4 quizzes, each having variance of 30
- The variance of your aggregate quiz score is 340
- The pairwise correlation between all quizzes is identical
- What is the correlation between quizzes (e.g. $cor(quiz1, quiz2)$)?

Summary of Expectation rules

- General formula: $E[X] = \mu_X = \sum_x xp(x)$
- Linearity: $E[aX + bY] = aE[X] + bE[Y]$
- Variance: $var(x) = \sigma_X^2 = E[X^2] - E[X]^2$
- Variance Linear Combinations:
  $var(aX + bY) = a^2 var(X) + b^2 var(Y) + 2ab\ cov(X, Y)$
  - $cov(X, Y) = \sigma_{XY} = E[XY] - E[X]E[Y]$
  - If Independent, $cov(X, Y) = 0$
- $\rho_{xy} = \frac{cov(X,Y)}{\sigma_x \sigma_y}$

CDF

- The pmf gives us all of the information we need to calculate probabilities, but what if we need to calculate $P(3 \leq X \leq 7)$?
  - Need $p(3) + p(4) + p(5) + p(6) + p(7)$
- One convenient tool is the cumulative distribution function (CDF): $F(x) = P(X \leq x)$
  - The above is then $F(7) - F(2)$

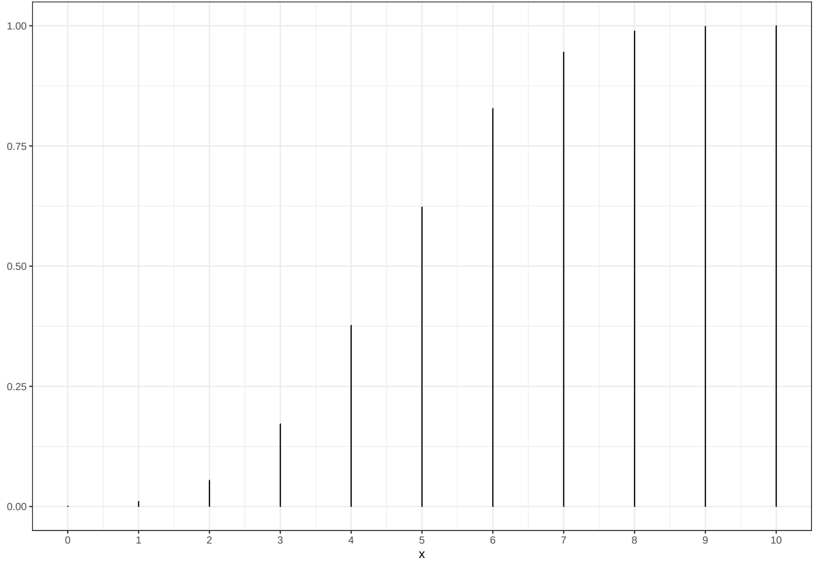
Uniform CDF

- The continuous uniform distribution has CDF $F(X) = x, 0 \leq x \leq 1$
- Calculate $P(.3 \leq x \leq .7)$

CDF

- The geometric distribution gives us the time until the first success for an event with independent draws of probability p
- The geometric distribution has CDF $P(X \leq x) = 1 - (1-p)^x, x \geq 1$
- How would we get the median?

CDF



Continuous random variables

- We previously had a uniform distribution defined by $F(x) = x, 0 \leq x \leq 1$
- Calculate $P(0.499 \leq x \leq 0.501)$
- What is $P(X = 0.5)$?

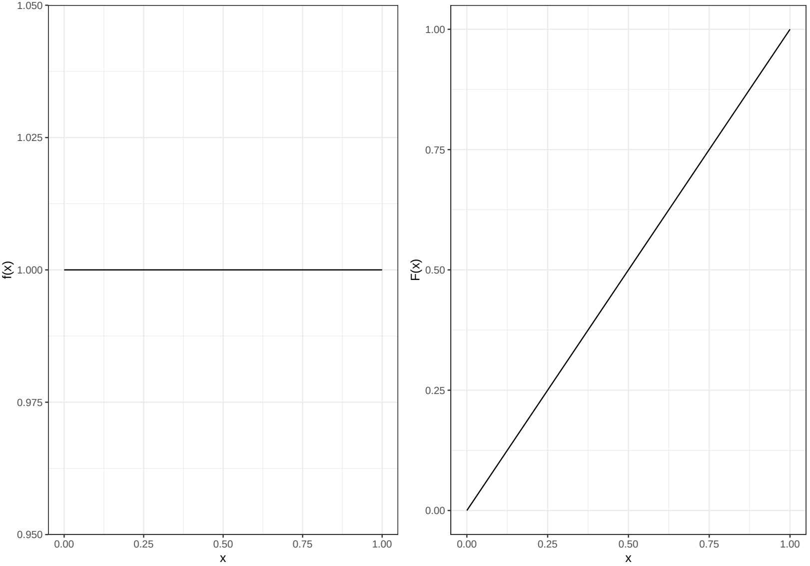
Continuous Random Variables

- The uniform distribution is tricker than it appears. We can calculate $P(X = x) = F(x) - F(x^-) = 0$ for every possible value of x
- There is 0 probability that any event happens. We say that this **almost certainly** does not happen. A probability of 0 does not mean impossible!
- The trick is that there are an uncountably infinite number of points between 0 and 1, each of which is equally likely.

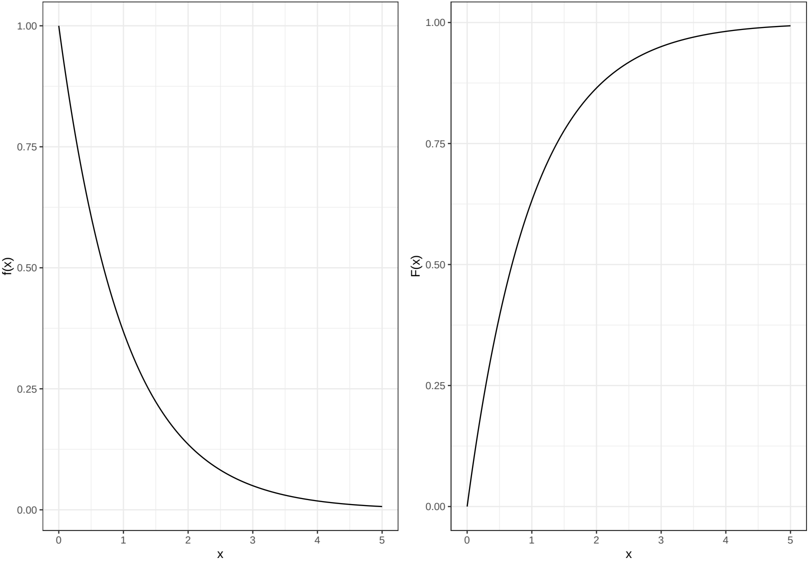
Continuous Random Variables

- We need calculus to add these up
  - $F(x) = \int_{-\infty}^{x} f(y)dy$ rather than $F(x) = \sum_{k=-\infty}^{x} p(k)$
- We call $f(x)$ the probability **density** function rather than the probability mass function
  - The values integrate to 1 rather than add to 1
  - These are not probabilities! They are **likelihoods**. They only give relative probabilities of occurring.

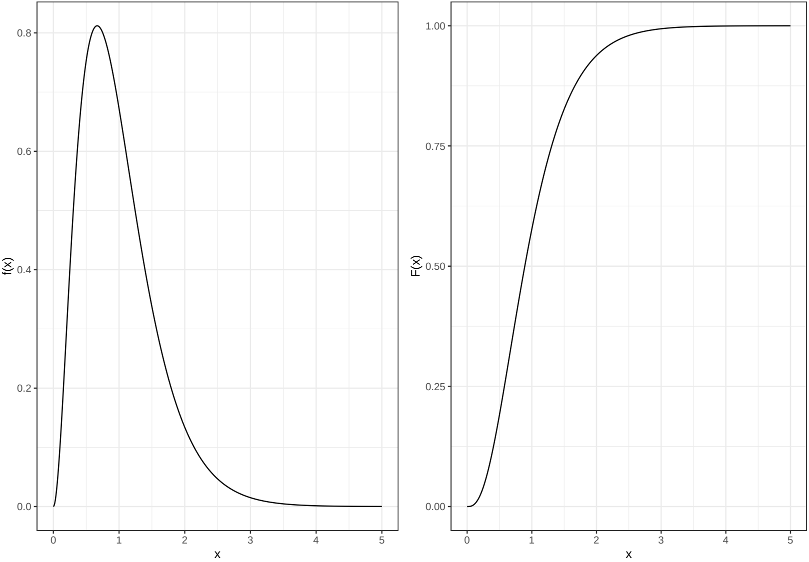
Continuous Random Variables: Uniform Graph



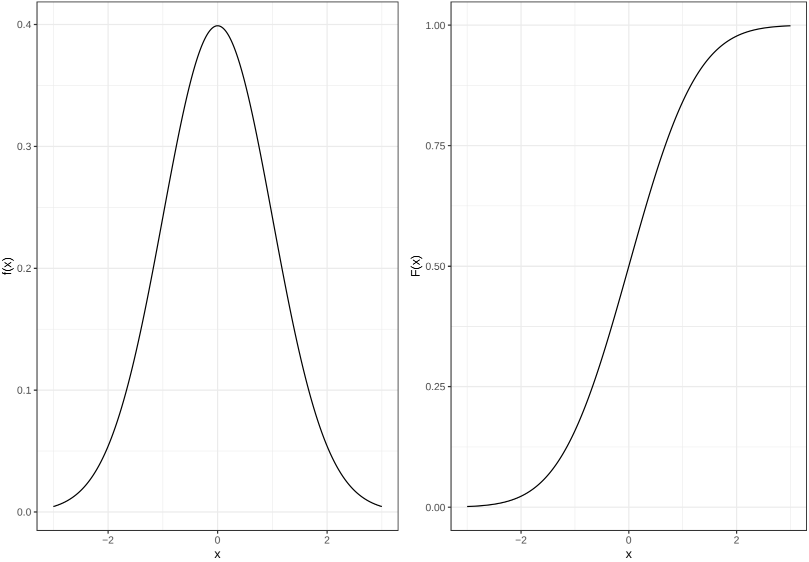
Continuous Random Variables: Exponential Graph



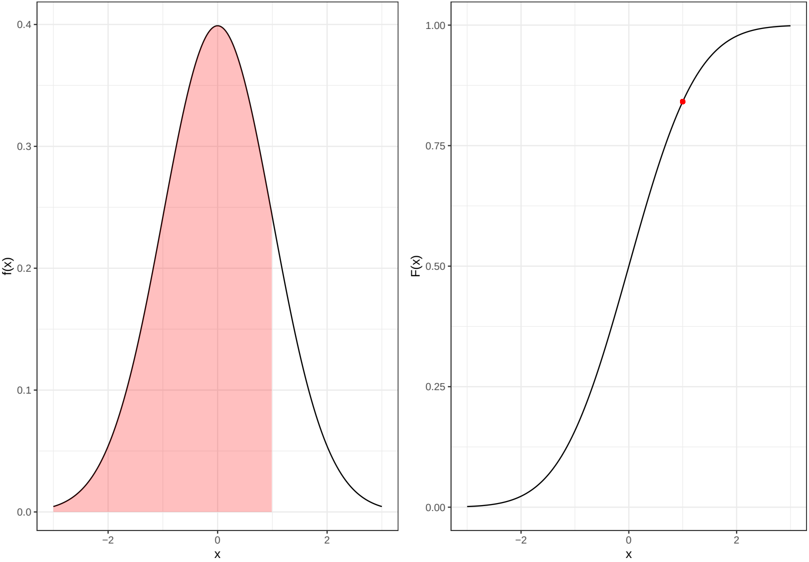
Continuous Random Variables: Gamma Graph



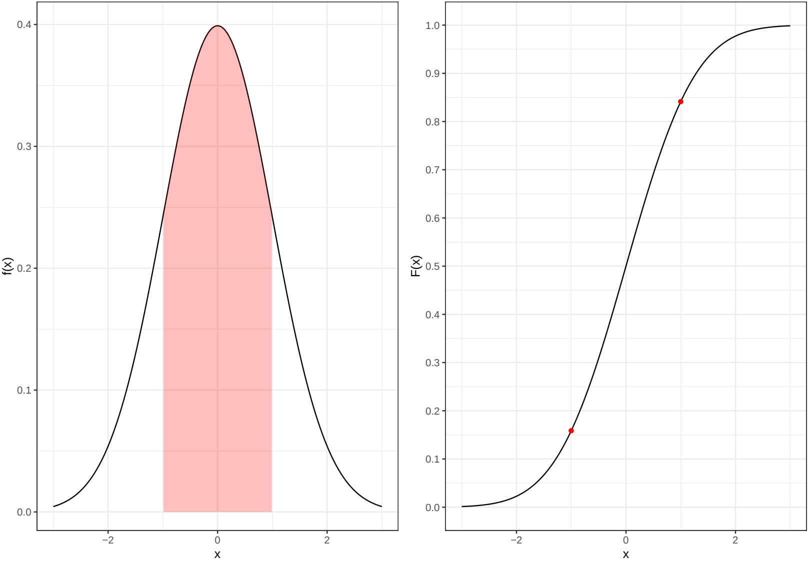
Continuous Random Variables: Normal Graph



Normal Probability



Normal Probability



The (standard) normal distribution

- The standard normal distribution has a distinct bell shape
- For now the formula just comes out of left field:
  - $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
  - Some neat integration tricks confirms $\int_{-\infty}^{\infty} f(x)dx = 1$
- $F(x) = \int_{-\infty}^{x} f(x)dx = \Phi(x)$
  - No closed form solution. We look up the values in a table or using software.
- $\mu = 0$ and $\sigma^2 = 1$

Calculating Standard Normal Probabilities

$P(-0.5 \leq x \leq 1)$

```
      x F(x)          x F(x)
1: -3.0 0.00        0.0 0.50
2: -2.5 0.01        0.5 0.69
3: -2.0 0.02        1.0 0.84
4: -1.5 0.07        1.5 0.93
5: -1.0 0.16        2.0 0.98
6: -0.5 0.31        2.5 0.99
7:  0.0 0.50        3.0 1.00
```

Calculating Standard Normal Probabilities

- $P(-.25 \leq x \leq 1.25)$

```
      x F(x)          x F(x)
1: 0.00 0.50       1.00 0.84
2: 0.25 0.60       1.25 0.89
3: 0.50 0.69       1.50 0.93
4: 0.75 0.77       1.75 0.96
5: 1.00 0.84       2.00 0.98
```

Quantiles

- What is the 99th percentile of a standard normal distribution?

```
     x  F(x)        x  F(x)
1: 2.0 0.977      2.5 0.994
2: 2.1 0.982      2.6 0.995
3: 2.2 0.986      2.7 0.997
4: 2.3 0.989      2.8 0.997
5: 2.4 0.992      2.9 0.998
6: 2.5 0.994      3.0 0.999
```

The Normal Family

- In general, we can refer to a normal distribution with mean μ and standard deviation σ
  - $X \sim N(\mu, \sigma)$
- The normal distribution has an important property: if $X \sim N(\mu, \sigma)$ and $Z = \frac{X-\mu}{\sigma}$, then $Z \sim N(0, 1)$
  - The normal distribution is the only finite-variance distribution with this property!
- This means that we only need the standard normal CDF
  - If we calculate the z-score of a normal random variable, we will have a standard normal distribution

Normal Calculations

- The height for males (in inches) in the US is approximately normally distributed with mean 69 and standard deviation 3
- Calculate the probability that a randomly selected male is between 66 and 72 inches (5'6" and 6')

Normal Approximation

- The normal distribution was independently discovered several times
  - If X is a Poisson random variable with mean λ, as $\lambda \to \infty$, $X \to N(\lambda, \sqrt{\lambda})$
  - If $X \sim Binom(n, p)$, $X \to N(\mu_X, \sigma_X^2)$ as $n \to \infty$
  - If $X \sim Gamma(\alpha, \lambda)$ $X \to N(\frac{\alpha}{\lambda}, \frac{\sqrt{\alpha}}{\lambda})$ as $r \to \infty$
  - If X is any finite variance distribution, and X_i represents independent samples from X, then $\bar{X} \sim N(\mu_x, \frac{\sigma_x}{\sqrt{n}})$ as $n \to \infty$

Normal Approximation

- Approximately normal distributions arise fairly frequently in nature, and their probabilities are easier to calculate than most other types of distributions
- We only need to know μ and σ^2 to generate probabilities using a normal random variable
- Suppose we flip a fair coin independently 1024 times. What is the probability that the we obtain between 480 and 544 heads?

Normal approximation

- Suppose we flip a fair coin independently 1024 times. What is the probability that the we obtain between 480 and 544 heads?

```
      x F(x)          x F(x)
1: -3.0 0.00        0.0 0.50
2: -2.5 0.01        0.5 0.69
3: -2.0 0.02        1.0 0.84
4: -1.5 0.07        1.5 0.93
5: -1.0 0.16        2.0 0.98
6: -0.5 0.31        2.5 0.99
7:  0.0 0.50        3.0 1.00
```

Normal Approximation

- Suppose we flip a fair coin independently 1024 times
- We become suspicious that the coin is biased if there is a less than 1 percent chance of observing at least as many heads as are obtained
- Calculate the fewest number of heads we would need to observe to become suspicious

Normal approximation

- Suppose we flip a fair coin independently 1024 times. What is the 99th percentile of heads we obtain?

```
       p x:p=F(x)
1: 0.900    1.282
2: 0.950    1.645
3: 0.975    1.960
4: 0.990    2.326
5: 0.995    2.576
```

Normal approximation

- Suppose we aren't concerned specifically that the coin is biased towards heads, but that the coin could be biased towards either heads or tails
- We become suspicious that the coin is biased if we observe a result outside of the middle 99 percent
- How many heads or tails would we need to observe in order to become suspicious that the coin is biased?

Normal approximation

- Suppose we flip a fair coin independently 1024 times. How many heads would we need to obtain to become suspicious of the coin?

```
       p x:p=F(x)
1: 0.900    1.282
2: 0.950    1.645
3: 0.975    1.960
4: 0.990    2.326
5: 0.995    2.576
```

A note on hypothesis testing

- Here we become suspicious of the coin if the **conditional probability** of obtaining a result this extreme is under 1 percent **given that the coin is fair**
- A coin that is slightly biased may appear to be fair
  - What is the probability that we are **not** suspicious of a biased coin that has $p = .51$?
  - This is a **false negative**

Calculating false negative for p=.51

- If our coin is biased with p=.51, we want to know the probability that we obtain at least 553 heads or less than 470 heads
- $\mu = .51 * 1024 = 522, \sigma = \sqrt{1024 * .51 * .49} = 16$
- $z_1 = -3.25, z_2 = 1.9375$
- $\Phi(-3.25) + (1 - \Phi(1.9375)) = 2.7\%$
  - We are very unlikely to suspect that this coin is biased!

A note on hypothesis testing

- Similarly, a coin that is truly unbiased will occasionally be suspicious
  - This is a **false positive**
  - We designed the probability of a false positive (conditional on a fair coin) to be 1%
- To calculate the probability of making an error, we need additional information

A note on hypothesis testing

- Suppose that there is a 50% chance that the coin is fair, and a 50% chance that the coin is biased with p=.51
- What is the probability that we misclassify the coin?
  - suspicious of fair coin: $.01 * .5 = .005$
  - not suspicous of biased coin: $.973 * .5 = .4865$
  - Probability of misclassification: $.005 + .4865 = .4915$