Math 3081 (Probability and Statistics)

Lecture #12 of 27 $\sim$ July 26th, 2021

The Central Limit Theorem and Applications

- The Central Limit Theorem
- The Normal Approximation to the Binomial Distribution
- Applications and Examples of the Central Limit Theorem
- Poisson Distributions

This material represents §2.3.2-2.3.3 from the course notes, and problems 5-11 from WeBWorK 4.

Recall, I

Last time, we introduced the normal distribution:

Definition

A random variable $N_{\mu,\sigma}$ is normally distributed with parameters μ and σ if its probability density function is $p_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$. The mean is μ and the standard deviation is σ.

- The standard normal distribution is $N_{0,1}$, with $\mu = 0$ and $\sigma = 1$.
- All normal distributions' pdfs are geometrically similar and have the famous "bell curve" shape.
- We have the "68-95.5-99.7" rule: these are the percentages of a normal distribution that will fall within 1, 2, and 3 standard deviations of the mean respectively.

Recall, II

We also introduced the central limit theorem to explain why the normal distribution often serves as a model:

Theorem (Central Limit Theorem)

Let X_1, X_2, ... , X_n be a sequence of independent, identically-distributed discrete or continuous random variables each with finite expected value μ and standard deviation $\sigma > 0$. Then the distribution of the random variable $Y_n = \dfrac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$ will approach the standard normal distribution (of mean 0 and standard deviation 1) as n tends to ∞: explicitly, for any real numbers $a \leq b$ we have $P(a \leq Y_n \leq b) \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$ as $n \to \infty$.

Central Limit Theorem, I

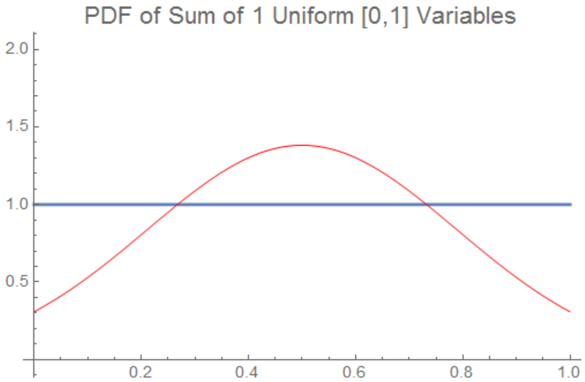
As mentioned last time, the central limit theorem is quite general.

- The idea is as follows: suppose we independently sample a random variable X a total of n times, to get values $X_1, X_2, \ldots, X_n$.
- Then $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = n\mu$, and also $\mathrm{var}(X_1 + X_2 + \cdots + X_n) = \mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n) = n\sigma^2$, so $\sigma(X_1 + \cdots + X_n) = \sqrt{n\sigma^2} = \sigma\sqrt{n}$.
- Thus, for $Y_n = \dfrac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$, by expected value properties we see that $E(Y_n) = 0$ and $\sigma(Y_n) = 1$.

The central limit theorem then says: if we take a sample size n tending to ∞, this normalized average Y_n approaches the standard normal distribution (expected value 0, standard deviation 1).

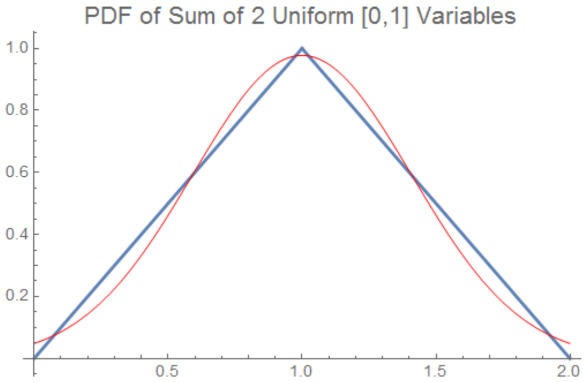
Central Limit Theorem, II

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$:



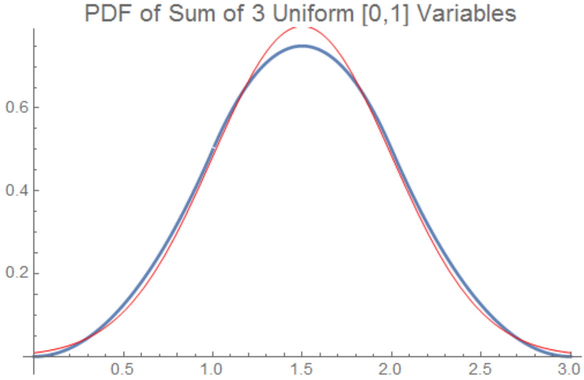
Central Limit Theorem, III

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue = sum of uniform variables, red = normal distribution):



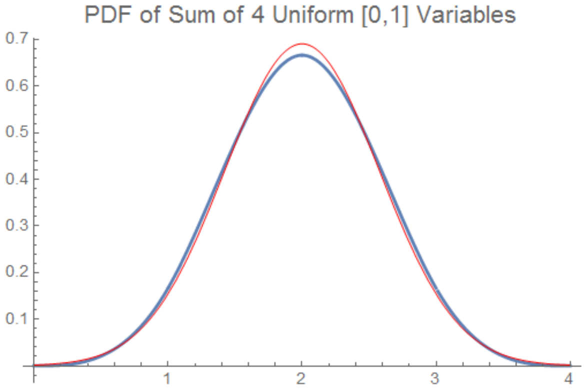
Central Limit Theorem, IV

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue = sum of uniform variables, red = normal distribution):



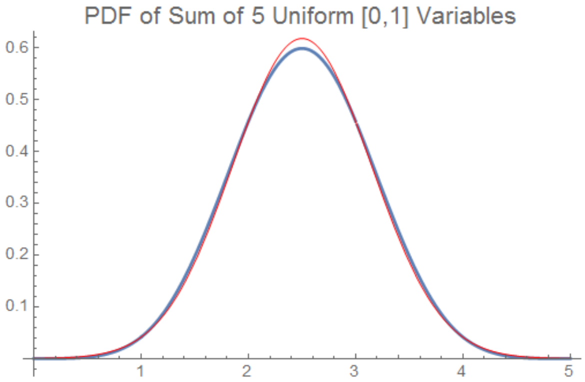
Central Limit Theorem, V

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue $=$ sum of uniform variables, red $=$ normal distribution):



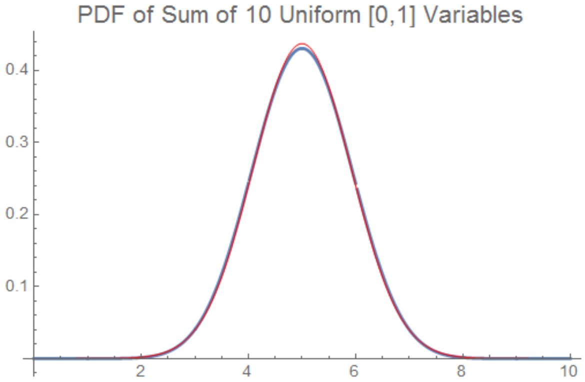
Central Limit Theorem, VI

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue = sum of uniform variables, red = normal distribution):



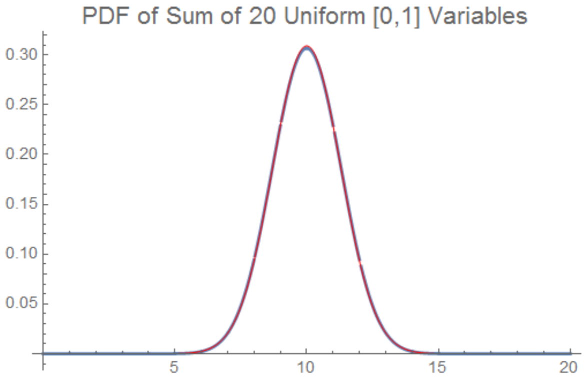
Central Limit Theorem, VII

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue = sum of uniform variables, red = normal distribution):



Central Limit Theorem, VIII

We can illustrate the central limit theorem quite convincingly for X the uniform distribution on $[0, 1]$ (blue $=$ sum of uniform variables, red $=$ normal distribution):



Central Limit Theorem, IX

As can be seen quite clearly from the graphs, the convergence of the sum distribution (hence also the normalized average) to a normal distribution is quite rapid.

- The central limit theorem says that the distribution of $Y_n = \dfrac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$ is approximately normal for large n.
- We can convert this back into a statement about the original sum, or alternatively, into a statement about the average $\frac{1}{n}(X_1 + X_2 + \cdots + X_n)$.
- Specifically, the central limit theorem says that both the sum and the average will be approximately normally distributed: the sum will have mean $n\mu$ and standard deviation $\sigma\sqrt{n}$, while the average will have mean μ and standard deviation $\sigma/\sqrt{n}$.

Normal Approximation to Binomial, I

As a first step, we apply this observation about sums when X is a Bernoulli random variable with success probability p.

- As noted on the previous slide, the central limit theorem implies that for n sufficiently large, the sum $X_1 + X_2 + \cdots + X_n$ will have a distribution that is approximately normal, with mean $n\mu$ and standard deviation $\sigma\sqrt{n}$.
- Since the Bernoulli random variable X has mean $\mu = p$ and standard deviation $\sigma = \sqrt{p(1-p)}$, that means $X_1 + X_2 + \cdots + X_n$ has an approximately normal distribution with mean np and standard deviation $\sqrt{np(1-p)}$.

Normal Approximation to Binomial, I

As a first step, we apply this observation about sums when X is a Bernoulli random variable with success probability p.

- As noted on the previous slide, the central limit theorem implies that for n sufficiently large, the sum $X_1 + X_2 + \cdots + X_n$ will have a distribution that is approximately normal, with mean $n\mu$ and standard deviation $\sigma\sqrt{n}$.
- Since the Bernoulli random variable X has mean $\mu = p$ and standard deviation $\sigma = \sqrt{p(1-p)}$, that means $X_1 + X_2 + \cdots + X_n$ has an approximately normal distribution with mean np and standard deviation $\sqrt{np(1-p)}$.
- But this distribution is simply the binomial distribution with parameters n and p!
- So what that means is: when n is large, we can approximate the binomial distribution with the normal distribution.

Normal Approximation to Binomial, II

We summarize the previous discussion as follows:

Approximation (Normal Approximation to Binomial Distribution)

If $0 < p < 1$ *and* n *is sufficiently large, then the binomial distribution with* n *trials and success probability* p *is well approximated by the normal distribution with the same mean* np *and standard deviation* $\sqrt{np(1-p)}$.

Normal Approximation to Binomial, II

We summarize the previous discussion as follows:

Approximation (Normal Approximation to Binomial Distribution)

If $0 < p < 1$ *and* n *is sufficiently large, then the binomial distribution with* n *trials and success probability* p *is well approximated by the normal distribution with the same mean* np *and standard deviation* $\sqrt{np(1-p)}$.

- Now, this is merely an approximation, so it is not really a theorem of any kind.
- It is primarily a heuristic estimate, and (historically speaking) was actually the first application of the normal distribution.

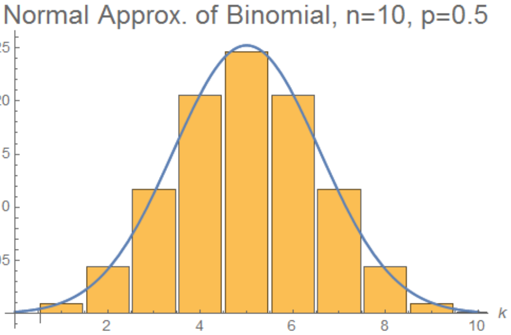
Normal Approximation to Binomial, III

As a practical matter, there are various heuristics that have been given to decide when the normal approximation to the binomial distribution falls into the category of “very good”.

- One “rule” states that the approximation will be good when np and $n(1-p)$ are both at least 5, and it increases in accuracy when np and $n(1-p)$ are larger.
- Another “rule” is that the approximation will be good when n is bigger than both $5p/(1-p)$ and $5(1-p)/p$.

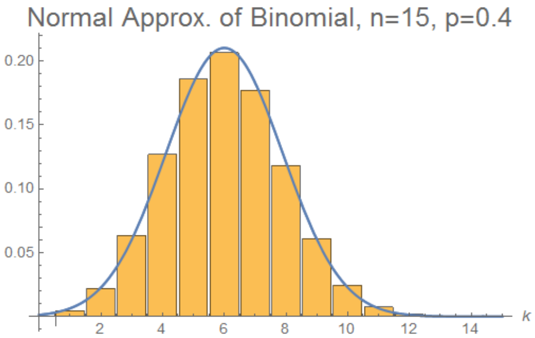
Normal Approximation to Binomial, IV

Here are some comparisons of binomial distributions (with various parameters n and p) with their corresponding normal approximations:



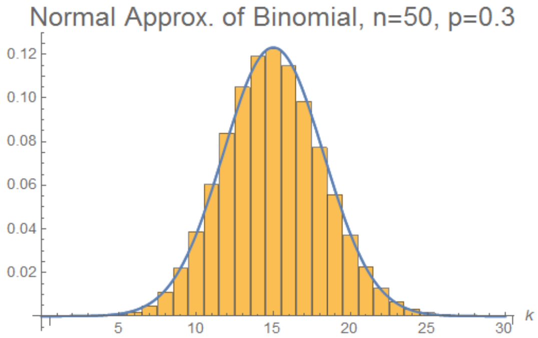
Normal Approximation to Binomial, V

Here are some comparisons of binomial distributions (with various parameters n and p) with their corresponding normal approximations:



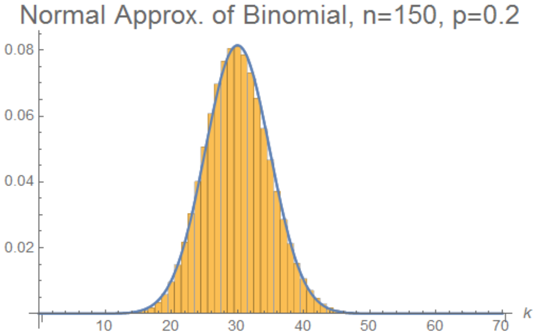
Normal Approximation to Binomial, VI

Here are some comparisons of binomial distributions (with various parameters n and p) with their corresponding normal approximations:



Normal Approximation to Binomial, VII

Here are some comparisons of binomial distributions (with various parameters n and p) with their corresponding normal approximations:



Normal Approximation to Binomial, VIII

In order to do explicit calculations, we must also adjust for the fact that the binomial distribution is discrete whereas the normal distribution is continuous.

- This type of adjustment is called a continuity correction: we approximate the probability $P(B = k)$ that the binomial random variable equals k with the probability $P(k - \frac{1}{2} \leq N \leq k + \frac{1}{2})$ that the normal random variable lands in the interval $[k - \frac{1}{2}, k + \frac{1}{2}]$.
- For an interval, we estimate $P(a \leq B \leq b)$ by $P(a - \frac{1}{2} \leq N \leq b + \frac{1}{2})$.
- We make a similar continuity correction whenever we approximate a discrete distribution by a continuous one, because in all such cases we need to compare areas to areas.

Normal Approximation to Binomial, IX

We can give a quick example for why the continuity correction is necessary.

- Consider the case of $n = 100$, $p = 1/2$, and computing $P(B = 50)$.
- If we did not make a continuity correction, we would estimate this value by $P(50 \leq N \leq 50)$, which is zero.
- But the actual value $P(B = 50) = \binom{100}{50}/2^{100} \approx 0.07959$, which is very far away from zero.
- On the other hand, since N has mean $np = 50$ and standard deviation $\sqrt{np(1-p)} = 5$, using the continuity correction gives an estimate $P(49.5 \leq N_{50,5} \leq 50.5) \approx 0.07966$, which is very accurate.

Normal Approximation to Binomial, X

Example: A fair coin is flipped 400 times.

1. Find the exact probability that the coin will land heads exactly 200 times, and compare to the result of the normal approximation (with continuity correction).
2. Repeat for the probability that the coin lands heads between 203 and 208 times inclusive.

Normal Approximation to Binomial, X

Example: A fair coin is flipped 400 times.

1. Find the exact probability that the coin will land heads exactly 200 times, and compare to the result of the normal approximation (with continuity correction).
2. Repeat for the probability that the coin lands heads between 203 and 208 times inclusive.

- The exact probabilities are binomially distributed, so we can simply write down formulas in terms of binomial coefficients. (Of course, we need a computer to calculate them.)
- The approximations also require values from the normal distribution, but these we could look up from the table.
- Here, we have $n = 400$ and $p = 1/2$, so the normal approximation to this binomial distribution has expected value $\mu = np = 200$ and standard deviation $\sigma = \sqrt{np(1-p)} = 10$.

Normal Approximation to Binomial, XI

Example: A fair coin is flipped 400 times.

1. Find the exact probability that the coin will land heads exactly 200 times, and compare to the result of the normal approximation (with continuity correction).

Normal Approximation to Binomial, XI

<u>Example</u>: A fair coin is flipped 400 times.

1. Find the exact probability that the coin will land heads exactly 200 times, and compare to the result of the normal approximation (with continuity correction).

- The number of heads is binomially distributed with $n = 400$ and $p = 1/2$, so the probability is $\binom{400}{200}/2^{400} \approx 3.99\%$.
- For the normal approximation, we have $\mu = np = 200$ and $\sigma = \sqrt{np(1-p)} = 10$.
- Thus, we want $P(199.5 \leq N_{200,10} \leq 200.5)$, which (if we use a table) is also $P(-0.05 \leq N_{0,1} \leq 0.05)$.
- Using a table or computer, we can find $P(N_{0,1} \leq -0.05) = 0.48006$ and $P(N_{0,05} \leq 0.1) = 0.51994$, so the estimate is $0.51994 - 0.48006 = 0.03988 \approx 3.99\%$.

Normal Approximation to Binomial, XII

Example: A fair coin is flipped 400 times.

2. Find the exact probability that the coin will land heads between 203 and 208 times, and compare to the result of the normal approximation (with continuity correction).

Normal Approximation to Binomial, XII

Example: A fair coin is flipped 400 times.

2. Find the exact probability that the coin will land heads between 203 and 208 times, and compare to the result of the normal approximation (with continuity correction).

- The probability is $\frac{1}{2^{400}}\left[\binom{400}{203}+\binom{400}{204}+\cdots+\binom{400}{208}\right]\approx 20.36\%$.
- For the approximation, we need $P(202.5 \le N_{200,10} \le 208.5)$, which from our discussion of z-scores is also equal to $P(0.25 \le N_{0,1} \le 0.85)$.
- Using a table or a computer, we can find $P(N_{0,1} \le 0.25) = 0.59871$ and $P(N_{0,1} \le 0.85) = 0.80234$, so the desired probability is $0.80234 - 0.59871 = 0.20363 \approx 20.36\%$.

Normal Approximation to Binomial, XIII

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

1. The probability that he faults between 10 and 20 times (inclusive) during the match.
2. The probability that he faults at most 15 times during the match.
3. The probability that he faults at least 25 times during the match.

Normal Approximation to Binomial, XIII

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

1. The probability that he faults between 10 and 20 times (inclusive) during the match.
2. The probability that he faults at most 15 times during the match.
3. The probability that he faults at least 25 times during the match.

- We use the normal approximation to the binomial distribution. This is a reasonable approach, because with $n = 150$ and $p = 0.151$ we have $np = 22.65$ and $n(1-p) = 127.35$, so the approximation should be good.
- The mean is $\mu = np = 22.65$ and the standard deviation is $\sigma = \sqrt{np(1-p)} \approx 4.3852$.

Normal Approximation to Binomial, XIV

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

1. The probability that he faults between 10 and 20 times (inclusive) during the match.

Normal Approximation to Binomial, XIV

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

1. The probability that he faults between 10 and 20 times (inclusive) during the match.

- With $\mu = 22.65$ and $\sigma = 4.3852$, we want to find the probability $P(9.5 < N_{\mu,\sigma} < 20.5)$.
- Using the z-score approach, this is equivalently asking for $P(-2.9987 < N_{0,1} < -0.4903)$.
- Either via a table or with a computer, the probability estimate is $P(9.5 < N_{\mu,\sigma} < 20.5) \approx 0.3106$.

Remark: The exact probability is ≈ 0.3188.

Normal Approximation to Binomial, XV

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

2. The probability that he faults at most 15 times during the match.

Normal Approximation to Binomial, XV

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

2. The probability that he faults at most 15 times during the match.

- With $\mu = 22.65$ and $\sigma = 4.3852$, we want to find the probability $P(N_{\mu,\sigma} < 15.5)$.
- Using the z-score approach, this is equivalently asking for $P(N_{0,1} < -1.6305)$.
- Either via a table or with a computer, the probability estimate is $P(N_{\mu,\sigma} < 15.5) \approx 0.0515$.

Remark: The exact probability is ≈ 0.0461.

Normal Approximation to Binomial, XVI

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

3. The probability that he faults at least 25 times during the match.

Normal Approximation to Binomial, XVI

Example: A tennis player serves 150 times during a match, and on an average serve he will fault 15.1% of the time, independently of any other serve. Estimate the following:

3. The probability that he faults at least 25 times during the match.

- With $\mu = 22.65$ and $\sigma = 4.3852$, we want to find the probability $P(N_{\mu,\sigma} > 24.5)$.
- Using the z-score approach, this is equivalently asking for $P(N_{0,1} > 0.4219)$.
- Either via a table or with a computer, the probability estimate is $P(N_{\mu,\sigma} > 24.5) \approx 0.3366$.

Remark: The exact probability is ≈ 0.3287.

Normal Stability, I

Another fundamental property of the normal distribution is that it is stable, in the sense that the sum of any number of independent normal distributions is also a normal distribution:

Proposition (Stability of Normal Distribution)

If $X_1, X_2, \ldots, X_n$ are independent normally-distributed random variables with means $\mu_1, \ldots, \mu_n$ and standard deviations $\sigma_1, \ldots, \sigma_n$, then the sum $X_1 + X_2 + \cdots + X_n$ is also normally distributed with mean $\mu_1 + \cdots + \mu_n$ and standard deviation $\sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$.

Normal Stability, II

Proof:

- It is a moderately straightforward calculation using the joint probability distribution to show that the distribution of the sum of two normally-distributed variables is also normally distributed. (Alternatively, this can be derived from the central limit theorem.)
- Thus, $X_1 + X_2$ is normally-distributed, hence so is $(X_1 + X_2) + X_3$, ... , and hence so is $X_1 + X_2 + \cdots + X_n$.
- For the mean, we have $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = \mu_1 + \mu_2 + \cdots + \mu_n$.
- For the standard deviation, because $X_1, X_2, \ldots, X_n$ are independent, we have $\mathrm{var}(X_1 + X_2 + \cdots + X_n) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \cdots + \mathrm{var}(X_n) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.
- So, $\sigma(X_1 + X_2 + \cdots + X_n) = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}$.

How To Use The Central Limit Theorem

The normal approximation to the binomial distribution was one of the first applications of the normal distribution and the central limit theorem, but there are lots of other applications as well.

- If we can write down the mean and standard deviation, then we can use the central limit theorem to estimate probabilities for arbitrary distributions obtained from repeated independent sampling, not just the binomial distribution.
- As we have also noted previously, the normal distribution is also stable under translation and rescaling: if X is normally distributed, then so is $aX + b$ for any fixed constants a and b.
- We can use these properties to analyze random variables that are obtained by adding, subtracting, or averaging independent (approximately) normal distributions, since the results will then also be (approximately) normally distributed.

More Examples, I

Example: Estimate the probability that if 420 fair dice are rolled, the total sum of all 420 rolls will be between 1460 and 1501 inclusive. (Note that if X is the random variable giving the result of one die roll, then $\mu_X = 7/2$ and $\sigma_X = \sqrt{35/12}$.)

More Examples, I

Example: Estimate the probability that if 420 fair dice are rolled, the total sum of all 420 rolls will be between 1460 and 1501 inclusive. (Note that if X is the random variable giving the result of one die roll, then $\mu_X = 7/2$ and $\sigma_X = \sqrt{35/12}$.)

- Since we are summing the results of 420 independent samplings of X, the central limit theorem tells us that the overall distribution of the sum will be closely approximated by a normal distribution with mean $420\mu_X = 1470$ and standard deviation $\sqrt{420}\mu_X = 35$.
- The desired probability is then $P(1459.5 \leq N_{1470,35} \leq 1501.5) = P(-0.3 \leq N_{0,1} \leq 0.9) \approx 0.4339$.

More Examples, II

Example: During a tournament, a professional poker player plays 2000 individual 5-card hands. Estimate the probability that they get at least 50 three-of-a-kind hands. (The probability of such a hand is 88/4165.)

More Examples, II

<u>Example</u>: During a tournament, a professional poker player plays 2000 individual 5-card hands. Estimate the probability that they get at least 50 three-of-a-kind hands. (The probability of such a hand is 88/4165.)

- The exact distribution of the number of three-of-a-kind hands will be binomial, with $n = 2000$ and $p = 88/4165$.
- We could write down a formula for the exact probability using the binomial distribution: it is $\binom{2000}{51} p^{51}(1-p)^{1949} + \cdots + \binom{2000}{2000} p^{2000}(1-p)^0$.
- This is hard to evaluate, but we can use the normal approximation: the mean is $\mu = np \approx 42.257$ and the standard deviation is $\sigma = \sqrt{np(1-p)} \approx 6.431$.
- Using the continuity correction, we see that the desired estimate is $P(N_{\mu,\sigma} > 49.5) \approx 0.1300$.

More Examples, IV

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

1. The expected mean and standard deviation of the average exam score in each class.
2. The probability that the average in class A is at least 81.
3. The probability that the average in class B is less than 79.
4. The distribution of the difference of the class averages.
5. The probability that class A outscores B on average by $\geq$ 1pt.
6. The probability that the class averages are within 0.2 points.

More Examples, IV

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

1. The expected mean and standard deviation of the average exam score in each class.
2. The probability that the average in class A is at least 81.
3. The probability that the average in class B is less than 79.
4. The distribution of the difference of the class averages.
5. The probability that class A outscores B on average by $\geq$ 1pt.
6. The probability that the class averages are within 0.2 points.

- Note that each class's score is normally distributed, since they are averages of independent, normally-distributed randomly variables.

More Examples, V

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

1. The expected mean and standard deviation of the average exam score in each class.

More Examples, V

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

1. The expected mean and standard deviation of the average exam score in each class.

- As we have previously noted, if a normally-distributed random variable with mean μ and standard deviation σ is sampled n times, then the mean of the average will be normally distributed with mean μ and standard deviation $\sigma/\sqrt{n}$.
- Thus, the average score in class A is normally distributed with mean 80 and standard deviation $6/\sqrt{9} = 2$.
- The average score in class B is normally distributed with mean 80 and standard deviation $6/\sqrt{16} = 1.5$.

More Examples, VI

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

2. The probability that the average in class A is at least 81.

More Examples, VI

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

2. The probability that the average in class A is at least 81.

- The average score in class A is normally distributed with mean 80 and standard deviation $6/\sqrt{9} = 2$.
- Thus, we want $P(N_{80,2} \geq 81) = P(N_{0,1} \geq 0.5) = 0.3085$.

3. The probability that the average in class B is less than 79.

More Examples, VI

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

2. The probability that the average in class A is at least 81.

- The average score in class A is normally distributed with mean 80 and standard deviation $6/\sqrt{9} = 2$.
- Thus, we want $P(N_{80,2} \geq 81) = P(N_{0,1} \geq 0.5) = 0.3085$.

3. The probability that the average in class B is less than 79.

- The average score in class B is normally distributed with mean 80 and standard deviation $6/\sqrt{16} = 1.5$.
- Thus, we want $P(N_{80,1.5} < 79) = P(N_{0,1} < -0.3333) = 0.2525$.

More Examples, VII

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

4. The distribution of the difference of the class averages.

More Examples, VII

<u>Example</u>: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

4. The distribution of the difference of the class averages.

- The average score in class A is normally distributed with mean 80 and standard deviation $6/\sqrt{9} = 2$, while the average score in class B is normally distributed with mean 80 and standard deviation $6/\sqrt{16} = 1.5$.
- The idea is to recognize that $A - B = A + (-B)$ and that $-B$ is also normally distributed (now with mean -80 and standard deviation 1.5).
- Thus, by our results, the random variable $A + (-B)$ will also be normally distributed with mean $80 + (-80) = 0$ and standard deviation $\sqrt{2^2 + 1.5^2} = 2.5$.

More Examples, VIII

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

5. The probability that class A outscores B on average by $\geq$ 1pt.

More Examples, VIII

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

5. The probability that class A outscores B on average by $\geq$ 1pt.

- Since $A - B$ is normally distributed with mean 0 and standard deviation 2.5, the desired probability is $P(N_{0,2.5} \geq 1) = P(N_{0,1} \geq 0.4) = 0.3446$.

6. The probability that the class averages are within 0.2 points.

More Examples, VIII

Example: A statistics instructor has two classes: Class A has 9 students while Class B has 16. The student scores are normally distributed with mean 80 and standard deviation 6. Find

5. The probability that class A outscores B on average by $\geq$ 1pt.

- Since $A - B$ is normally distributed with mean 0 and standard deviation 2.5, the desired probability is $P(N_{0,2.5} \geq 1) = P(N_{0,1} \geq 0.4) = 0.3446$.

6. The probability that the class averages are within 0.2 points.

- Since $A - B$ is normally distributed with mean 0 and standard deviation 2.5, the desired probability is $P(-0.2 \leq N_{0,2.5} \leq 0.2) = P(-0.08 \leq N_{0,1} \leq 0.08) = 0.0638$.

More Examples, IX

The ideas in this last example form the basis for many approaches in statistical testing, since these calculations give a way of determining how likely it is that a difference in sampling averages has occurred by chance, if the means of the distributions were actually equal. (We will discuss this much more in chapter 4.)

We will also use similar ideas in the next chapter when we discuss confidence intervals derived from normal distributions.

More Examples, X

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

1. Find the expected number of points scored.
2. Find the standard deviation in the number of points scored.
3. Estimate the probability that the team scores ≥ 110 points.
4. Estimate the probability that the team scores 95 points.

More Examples, X

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

1. Find the expected number of points scored.
2. Find the standard deviation in the number of points scored.
3. Estimate the probability that the team scores ≥ 110 points.
4. Estimate the probability that the team scores 95 points.

- The total number of free throws, two-pointers, and three-pointers will each be binomially distributed.
- Since the values of np and $n(1-p)$ are fairly large for each of these three distributions, they will be well approximated by the corresponding normal distributions with the same mean and standard deviation. The total number of points will then be a weighted sum of these approximately normal distributions, hence will also be approximately normal.

More Examples, XI

<u>Example</u>: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

1. Find the expected number of points scored.

More Examples, XI

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

1. Find the expected number of points scored.

- The number of free throws has $n = 15$ and $p = 0.75$, so the expected number is $15 \cdot 0.75 = 11.25$.
- The number of two-pointers has $n = 60$ and $p = 0.50$ hence the expected number of two-pointers is $60 \cdot 0.50 = 30$. Since two-pointers are worth 2 points each, the expected number of points is $2 \cdot 30 = 60$.
- In the same way, the number of three-pointers has $n = 25$ and $p = 0.35$, so the expected number of points from three-pointers is $3 \cdot 25 \cdot 0.35 = 26.25$.
- Thus, the expected total number of points is $11.25 + 60 + 26.25 = 97.5$.

More Examples, XII

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

2. Find the standard deviation in the number of points scored.

More Examples, XII

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

2. Find the standard deviation in the number of points scored.

- For free throws ($n = 15$, $p = 0.75$), the standard deviation is $\sqrt{15 \cdot 0.75 \cdot 0.25} \approx 1.6771$.
- For two-pointers ($n = 60$, $p = 0.50$), the standard deviation is $\sqrt{60 \cdot 0.50 \cdot 0.50} \approx 3.8730$, so the standard deviation in the number of points is 2 times this, which is 7.7460.
- For three-pointers ($n = 25$, $p = 0.35$), the standard deviation in the number of points is $3\sqrt{25 \cdot 0.35 \cdot 0.65} = 7.1545$.
- Thus, the standard deviation of the total is $\sqrt{1.6771^2 + 7.7460^2 + 7.1545^2} \approx 10.6771$.

More Examples, XIII

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

3. Estimate the probability that the team scores ≥ 110 points.

More Examples, XIII

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

3. Estimate the probability that the team scores ≥ 110 points.

- We have just found that the mean number of points is 97.5 and the standard deviation is 10.6771.
- Since the distribution is approximately normal, the probability that the team will score at least 110 points is given by $P(N_{97.5,10.6771} \geq 110) = P(N_{0,1} \geq 1.1707) \approx 0.1209$, or roughly 12.1%.

More Examples, XIV

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

4. Estimate the probability that the team scores 95 points.

More Examples, XIV

Example: In a typical game, a basketball team attempts 15 free throws, 60 two-pointers, and 25 three-pointers, which independently score 75%, 50%, and 35% of the time respectively.

4. Estimate the probability that the team scores 95 points.

- To estimate this probability, we need to use a continuity correction to handle the fact that the total number of points is discrete rather than continuous. (In fact, we should also have done this in the previous calculation! It changes the probability estimate to 0.1305.)
- Here, we can compute $P(94.5 \leq N_{97.5,10.6771} \leq 95.5) = P(-0.2810 \leq N_{0,1} \leq -0.1873) \approx 0.0363$, or roughly 3.63%.

Poisson Distributions, I

We now introduce our next model, the Poisson distribution:

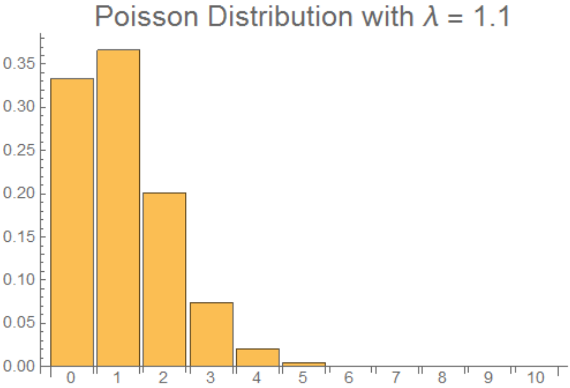
Definition

The <u>*Poisson distribution with parameter* $\lambda > 0$</u> *is the discrete random variable* X *that takes the nonnegative integer value* n *with probability* $P(X = n) = \dfrac{\lambda^n e^{-\lambda}}{n!}$.

- Note that this is in fact a valid probability distribution because $\displaystyle\sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} = e^{-\lambda} \cdot \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1$, where we used the Taylor series expansion $\displaystyle e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

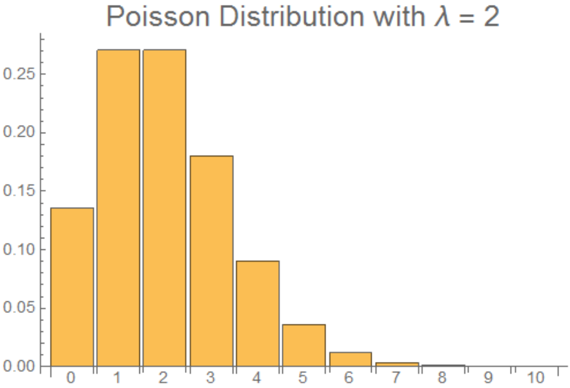
Poisson Distributions, II

Here are some plots of Poisson pdfs:



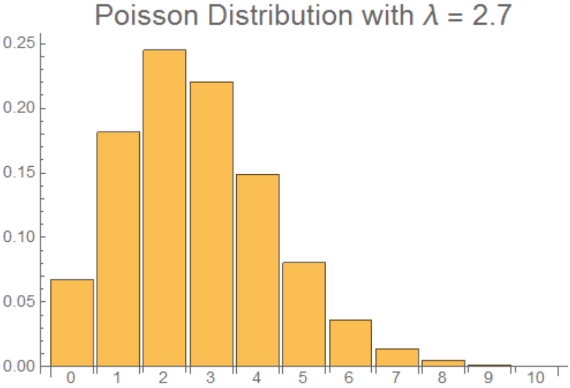
Poisson Distributions, III

Here are some plots of Poisson pdfs:



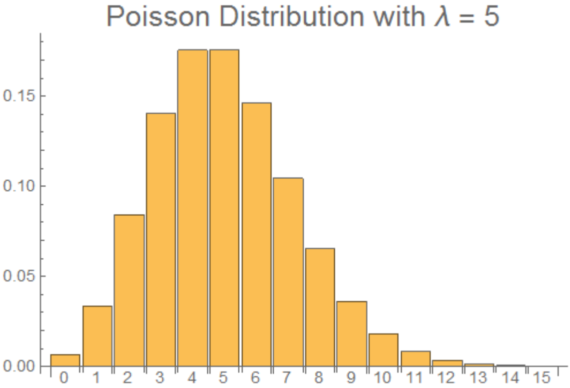
Poisson Distributions, IV

Here are some plots of Poisson pdfs:



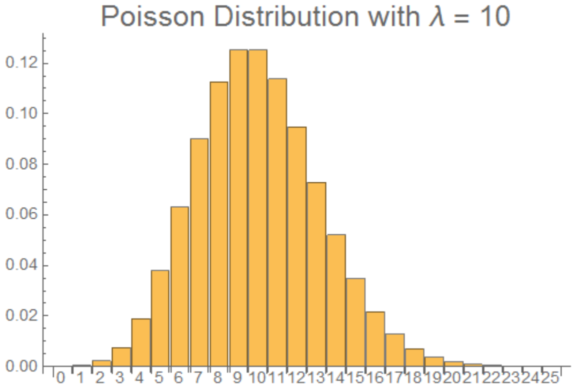
Poisson Distributions, V

Here are some plots of Poisson pdfs:



Poisson Distributions, VII

Here are some plots of Poisson pdfs:



Poisson Distributions, VIII

As might be suggested by the plots, the peak always occurs near $x = \lambda$:

- Since $P(X = n) = \dfrac{\lambda^n e^{-\lambda}}{n!}$, we see $\dfrac{P(X = n+1)}{P(X = n)} = \dfrac{\lambda}{n+1}$.
- Thus when $n < \lambda - 1$, $P(X = n+1) > P(X = n)$ while when $n > \lambda - 1$, $P(X = n+1) < P(X = n)$.
- This means the values $p(X = 1), p(X = 2), \ldots$ increase until λ exceeds $n - 1$, and then decrease. That means the peak occurs when n is the greatest integer less than or equal to λ (and when λ is exactly an integer, the peak is shared between $\lambda - 1$ and λ).

Poisson Distributions, IX

We can compute the expected value and variance:

- If X has a Poisson distribution with parameter λ, then $E(X) = \sum_{n=0}^{\infty} n\dfrac{\lambda^n e^{-\lambda}}{n!} = \lambda \sum_{n=1}^{\infty} \dfrac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} = \lambda,$
- Also, $E(X^2) = \sum_{n=0}^{\infty} n^2 \dfrac{\lambda^n e^{-\lambda}}{n!} =$ $\sum_{n=1}^{\infty} \lambda \dfrac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} + \sum_{n=2}^{\infty} \lambda^2 \dfrac{\lambda^{n-2} e^{-\lambda}}{(n-2)!} = \lambda + \lambda^2$, and so $\mathrm{var}(X) = E(X^2) - E(X)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda.$
- Thus, the expected value of a Poisson-distributed random variable is λ, and its variance is also λ.

Poisson Distributions, X

Example: If the random variable X has a Poisson distribution with $\lambda = 4$, find

1. $P(X = 2)$.
2. $P(X < 4)$.
3. $P(X \geq 3)$.
4. The expected value and standard deviation of X.
5. The probability that $X = 2$ given that $X > 1$.

Poisson Distributions, X

Example: If the random variable X has a Poisson distribution with $\lambda = 4$, find

1. $P(X = 2)$.
2. $P(X < 4)$.
3. $P(X \geq 3)$.
4. The expected value and standard deviation of X.
5. The probability that $X = 2$ given that $X > 1$.

- Here are the first few values of $P(X = n) = \dfrac{\lambda^n e^{-\lambda}}{n!}$:

n	0	1	2	3	4	5	6	7
$P(X = n)$	0.0183	0.0733	0.1465	0.1954	0.1954	0.1563	0.1042	0.0595

Poisson Distributions, XI

<u>Example</u>: If the random variable X has a Poisson distribution with $\lambda = 4$, find

1. $P(X = 2)$.

n	0	1	2	3	4	5	6	7
$P(X = n)$	0.0183	0.0733	0.1465	0.1954	0.1954	0.1563	0.1042	0.0595

- Using the table, $P(X = 2) \approx 0.1465$.

2. $P(X < 4)$.

- Using the table, $P(X < 4) = 0.0183 + 0.0733 + 0.1465 + 0.1954 \approx 0.4335$.

3. $P(X > 3)$.

- Using the table, $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.0183 - 0.0733 - 0.1465 - 0.1954 \approx 0.5665$.

Poisson Distributions, XII

Example: If the random variable X has a Poisson distribution with $\lambda = 4$, find

4. The expected value and standard deviation of X.

- The expected value is $\lambda = 4$ and the standard deviation is $\sqrt{\lambda} = 2$.

5. The probability that $X = 2$ given that $X > 1$.

n	0	1	2	3	4	5	6	7
$P(X = n)$	0.0183	0.0733	0.1465	0.1954	0.1954	0.1563	0.1042	0.0595

- This is $P(X = 2|X > 1) = P(X = 2 \cap X > 1)/P(X > 1) = P(X = 2)/P(X > 1)$.
- Since $P(X = 2) = 0.1954$ and $P(X > 1) = 1 - P(X \leq 1) = 1 - 0.0183 - 0.0733 = 0.9084$, the probability is $0.1465/0.9084 = 0.1613$.

Poisson Distributions, XIII

The Poisson distribution arises in the analysis of systems having a large number of independent events each of which occurs rarely.

- Specifically, suppose that we want to count the number X of times that a rare event occurs, under the hypothesis that occurrences of the rare event are independent.
- As we will discuss next time, the probability distribution of the number of rare independent events occurring in a fixed interval, under the assumption that the average number of events per interval is λ, will be Poisson with parameter λ.

Poisson Distributions, XIV

Examples (of quantities with a Poisson model):

- The number of soldiers killed by horse-kicks each year in the Prussian cavalry. (One of the first historical applications.)
- The number of calls received by a customer service center.
- The number of mutations created on a DNA strand during replication.
- The number of customers arriving at a restaurant or shop.
- The number of insurance claims during a given month.
- The number of earthquakes during a given month.
- The number of goals scored by a hockey or soccer team during a game.
- The number of decay events observed in a radioactive sample with a long half-life.
- The number of cases of a rare cancer in a given county.

Summary

We discussed the central limit theorem and the normal approximation to the binomial distribution.

We did some additional examples illustrating how the central limit theorem can be used to make estimates of probabilities.

We introduced the Poisson distribution.

Next lecture: Applications of the Poisson and exponential distributions.