Lecture 6.2, MATH-57091 Probability and Statistics for High-School Teachers.

Artem Zvavitch

Department of Mathematical Sciences, Kent State University

September, 29 - October 3, 2014.

Weak law of large numbers.

Let $X_1, X_2, \ldots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \ldots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Weak law of large numbers.

Let $X_1, X_2, \dots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality.

Weak law of large numbers.

Let $X_1, X_2, \dots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality. Indeed, the inequality tells us that if Y is a random variable with mean $\mathbb{E}Y$ and variance $Var(Y)$, then for all $k > 0$:

$$P(|Y - \mathbb{E}Y| \geq k) \leq \frac{Var(Y)}{k^2}.$$

Weak law of large numbers.

Let $X_1, X_2, \ldots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \ldots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality. Indeed, the inequality tells us that if Y is a random variable with mean $\mathbb{E}Y$ and variance $Var(Y)$, then for all $k > 0$:

$$P(|Y - \mathbb{E}Y| \geq k) \leq \frac{Var(Y)}{k^2}.$$

We apply the above inequality to $Y = \frac{X_1+X_2+\cdots+X_n}{n}$,

Weak law of large numbers.

Let $X_1, X_2, \dots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality. Indeed, the inequality tells us that if Y is a random variable with mean $\mathbb{E}Y$ and variance $Var(Y)$, then for all $k > 0$:

$$P(|Y - \mathbb{E}Y| \geq k) \leq \frac{Var(Y)}{k^2}.$$

We apply the above inequality to $Y = \frac{X_1+X_2+\dots+X_n}{n}$, then

$$\mathbb{E}Y = \mathbb{E}\frac{X_1 + X_2 + \dots + X_n}{n} = \frac{n\mu}{n} = \mu.$$

Weak law of large numbers.

Let $X_1, X_2, \dots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality. Indeed, the inequality tells us that if Y is a random variable with mean $\mathbb{E}Y$ and variance $Var(Y)$, then for all $k > 0$:

$$P(|Y - \mathbb{E}Y| \geq k) \leq \frac{Var(Y)}{k^2}.$$

We apply the above inequality to $Y = \frac{X_1+X_2+\dots+X_n}{n}$, then

$$\mathbb{E}Y = \mathbb{E}\frac{X_1 + X_2 + \dots + X_n}{n} = \frac{n\mu}{n} = \mu.$$

and (all X_i are independent!)

$$Var(Y) = Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n^2} Var\left(X_1 + X_2 + \dots + X_n\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Weak law of large numbers.

Let $X_1, X_2, \dots$ be independent random variables having the same mean and variance (i.e. $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$ for all $i = 1, 2, \dots$), then for all $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} = 0.$$

Proof: The main idea is to use the powerful and cool Chebychev's inequality. Indeed, the inequality tells us that if Y is a random variable with mean $\mathbb{E}Y$ and variance $Var(Y)$, then for all $k > 0$:

$$P(|Y - \mathbb{E}Y| \geq k) \leq \frac{Var(Y)}{k^2}.$$

We apply the above inequality to $Y = \frac{X_1+X_2+\cdots+X_n}{n}$, then

$$\mathbb{E}Y = \mathbb{E}\frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{n\mu}{n} = \mu.$$

and (all X_i are independent!)

$$Var(Y) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n^2} Var\left(X_1 + X_2 + \cdots + X_n\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Thus Chebychev's inequality gives us:

$$P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Strong law of large numbers.

The next theorem, known as **strong law of large numbers** is one of the most well-known results in probability theory. It stays that the average of a sequence of independent random variables having the same distribution will, with probability 1 converge to the mean of that distribution:

The next theorem, known as **strong law of large numbers** is one of the most well-known results in probability theory. It stays that the average of a sequence of independent random variables having the same distribution will, with probability 1 converge to the mean of that distribution:

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$. Then, with probability 1:

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mu, \text{ as } n \to \infty.$$

Example:

Consider a sequence of independent trials when each time we are trying for an event A to occur(say to see number 5 when we roll a die),

Example:

Consider a sequence of independent trials when each time we are trying for an event A to occur(say to see number 5 when we roll a die), let $P(A)$ be the probability that A occurs on any particular trial (so in example of number 5 - $P(A) = 1/6$).

Example:

Consider a sequence of independent trials when each time we are trying for an event A to occur(say to see number 5 when we roll a die), let $P(A)$ be the probability that A occurs on any particular trial (so in example of number 5 - $P(A) = 1/6$). As before, letting

$$X_i = \begin{cases} 1, & \text{if } A \text{ occurs on the } i\text{th trial,} \\ 0, & \text{if } A \text{ does not occur on the } i\text{th trial.} \end{cases}$$

Consider a sequence of independent trials when each time we are trying for an event A to occur(say to see number 5 when we roll a die), let $P(A)$ be the probability that A occurs on any particular trial (so in example of number 5 - $P(A) = 1/6$). As before, letting

$$X_i = \begin{cases} 1, & \text{if } A \text{ occurs on the } i\text{th trial,} \\ 0, & \text{if } A \text{ does not occur on the } i\text{th trial.} \end{cases}$$

Then we have by the strong law of large numbers that, with probability 1:

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mathbb{E}X_1 = P(A), \text{ as } n \to \infty.$$

Example:

Consider a sequence of independent trials when each time we are trying for an event A to occur(say to see number 5 when we roll a die), let $P(A)$ be the probability that A occurs on any particular trial (so in example of number 5 - $P(A) = 1/6$). As before, letting

$$X_i = \begin{cases} 1, & \text{if } A \text{ occurs on the } i\text{th trial}, \\ 0, & \text{if } A \text{ does not occur on the } i\text{th trial}. \end{cases}$$

Then we have by the strong law of large numbers that, with probability 1:

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mathbb{E}X_1 = P(A), \text{ as } n \to \infty.$$

Since $X_1 + X_2 + \cdots + X_n$ represents the number of times that even A occurs in the first n trials, we may interpret the above equation as stating that, with probability 1, the limiting proportion of time that the event A occurs is just $P(A)$.

Another result which makes probability super cool is **central limit theorem**.

Another result which makes probability super cool is **central limit theorem**. It is very important from the theoretical point of view, but, more over, it is essential practical tool. It gives us quite a simple method to approximate probabilities for sums of independent random variables.

Another result which makes probability super cool is **central limit theorem**. It is very important from the theoretical point of view, but, more over, it is essential practical tool. It gives us quite a simple method to approximate probabilities for sums of independent random variables. It is also explains why all of us (and especially people who do Statistics) love Normal Distribution: the empirical frequencies of so many natural "populations" exhibit a bell-shaped (NORMAL) curve.

Another result which makes probability super cool is **central limit theorem**. It is very important from the theoretical point of view, but, more over, it is essential practical tool. It gives us quite a simple method to approximate probabilities for sums of independent random variables. It is also explains why all of us (and especially people who do Statistics) love Normal Distribution: the empirical frequencies of so many natural "populations" exhibit a bell-shaped (NORMAL) curve.

Let $X_1, X_2, \dots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$.

Another result which makes probability super cool is **central limit theorem**. It is very important from the theoretical point of view, but, more over, it is essential practical tool. It gives us quite a simple method to approximate probabilities for sums of independent random variables. It is also explains why all of us (and especially people who do Statistics) love Normal Distribution: the empirical frequencies of so many natural "populations" exhibit a bell-shaped (NORMAL) curve.

Let $X_1, X_2, \dots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0, 1)$ as $n \to \infty$.

Another result which makes probability super cool is **central limit theorem**. It is very important from the theoretical point of view, but, more over, it is essential practical tool. It gives us quite a simple method to approximate probabilities for sums of independent random variables. It is also explains why all of us (and especially people who do Statistics) love Normal Distribution: the empirical frequencies of so many natural "populations" exhibit a bell-shaped (NORMAL) curve.

Let $X_1, X_2, \dots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}\, dx$$

as $n \to \infty$.

Strong law of large numbers.

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}\, dx$$

as $n \to \infty$.

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

as $n \to \infty$.

It also helps to note that

$$\mathbb{E}\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma\sqrt{n}}\mathbb{E}(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$$

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

as $n \to \infty$.

It also helps to note that

$$\mathbb{E}\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma\sqrt{n}} \mathbb{E}(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$$

and

$$Var\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma^2 n} Var(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma^2 n} Var(X_1 + X_2 + \cdots + X_n)$$

Let $X_1, X_2, \ldots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

as $n \to \infty$.

It also helps to note that

$$\mathbb{E}\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma\sqrt{n}}\mathbb{E}(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$$

and

$$Var\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma^2 n} Var(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma^2 n} Var(X_1 + X_2 + \cdots + X_n)$$

using independence we get

$$= \frac{1}{\sigma^2 n} n\sigma^2 = 1.$$

Let $X_1, X_2, \dots$ be a sequence of independent random variables having common distribution and let $\mathbb{E}X_i = \mu$ and $Var(X_i) = \sigma^2$. Then, the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal distribution $N(0,1)$ as $n \to \infty$. That is

$$P\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < a\right) \to P(N(0,1) < a) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{x^2}{2}}dx$$

as $n \to \infty$.

It also helps to note that

$$\mathbb{E}\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma\sqrt{n}}\mathbb{E}(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$$

and

$$Var\left(\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma^2 n}Var(X_1 + X_2 + \cdots + X_n - n\mu) = \frac{1}{\sigma^2 n}Var(X_1 + X_2 + \cdots + X_n)$$

using independence we get

$$= \frac{1}{\sigma^2 n} n\sigma^2 = 1.$$

Again, the amassing fact about this theorem (as well as about two previous) is that it works for very generic distribution of X_i is long as they are independent and identically distributed with bounded expectation and variance.

An idea of an application.

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables,

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables, then $\mathbb{E}X = n\mathbb{E}X_1$ and $Var(X) = nVar(X_1)$ and

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables, then $\mathbb{E}X = n\mathbb{E}X_1$ and $Var(X) = nVar(X_1)$ and

$$\frac{X_1 + X_2 + \cdots + X_n - n\mathbb{E}X_1}{\sqrt{Var(X_1)}\sqrt{n}} = \frac{X - \mathbb{E}X}{\sqrt{Var(X)}}.$$

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables, then $\mathbb{E}X = n\mathbb{E}X_1$ and $Var(X) = nVar(X_1)$ and

$$\frac{X_1 + X_2 + \cdots + X_n - n\mathbb{E}X_1}{\sqrt{Var(X_1)}\sqrt{n}} = \frac{X - \mathbb{E}X}{\sqrt{Var(X)}}.$$

If X is a Binomial random variable with parameters n and p, then X has the same distribution as the sum of n independent Bernoulli random variables each with parameter p thus

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables, then $\mathbb{E}X = n\mathbb{E}X_1$ and $Var(X) = nVar(X_1)$ and

$$\frac{X_1 + X_2 + \cdots + X_n - n\mathbb{E}X_1}{\sqrt{Var(X_1)}\sqrt{n}} = \frac{X - \mathbb{E}X}{\sqrt{Var(X)}}.$$

If X is a Binomial random variable with parameters n and p, then X has the same distribution as the sum of n independent Bernoulli random variables each with parameter p thus

$$\frac{X - \mathbb{E}X}{\sqrt{Var(X)}} = \frac{X - np}{\sqrt{np(1-p)}}.$$

An idea of an application.

Assume that $X = X_1 + X_2 + \cdots + X_n$ where all X_i independent identically distributed random variables, then $\mathbb{E}X = n\mathbb{E}X_1$ and $Var(X) = nVar(X_1)$ and

$$\frac{X_1 + X_2 + \cdots + X_n - n\mathbb{E}X_1}{\sqrt{Var(X_1)}\sqrt{n}} = \frac{X - \mathbb{E}X}{\sqrt{Var(X)}}.$$

If X is a Binomial random variable with parameters n and p, then X has the same distribution as the sum of n independent Bernoulli random variables each with parameter p thus

$$\frac{X - \mathbb{E}X}{\sqrt{Var(X)}} = \frac{X - np}{\sqrt{np(1-p)}}.$$

So $\frac{X-np}{\sqrt{np(1-p)}}$ approaches the standard normal distribution as n approaches ∞. It is known that the normal approximation will, in general, be reasonably good for values of n satisfying $np(1-p) \geq 10$.

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002.$

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002$.
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem.

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002$.
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem. We note that X is a descrete random variable and normal distribution is continuous one. We need a small "trick" to make them work together:

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15} \left(\frac{1}{6}\right)^{15} \left(\frac{5}{6}\right)^{85} \approx .01002.$
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem. We note that X is a descrete random variable and normal distribution is continuous one. We need a small "trick" to make them work together:

$$P(X = 15) = P(14.5 < X < 15.5) = P\left(\frac{14.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{15.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}}\right)$$

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002.$
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem. We note that X is a descrete random variable and normal distribution is continuous one. We need a small "trick" to make them work together:

$$P(X = 15) = P(14.5 < X < 15.5) = P\left(\frac{14.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{15.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}}\right)$$

$$= P\left(-.58 < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < -.31\right) \approx \Phi(-.31) - \Phi(-.58)$$

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002$.
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem. We note that X is a descrete random variable and normal distribution is continuous one. We need a small "trick" to make them work together:

$$P(X = 15) = P(14.5 < X < 15.5) = P\left(\frac{14.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{15.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}}\right)$$

$$= P\left(-.58 < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < -.31\right) \approx \Phi(-.31) - \Phi(-.58)$$

Where

$$\Phi(t) = P(N(0,1) < t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{y^2}{2}}\, dy.$$

Let X be the number of times that a die, flipped 100 times shows 6. Find the probability that $X = 15$. Use the normal approximation and then compare your it to the exact value.

Solution: We know that $P(X = 15) = \binom{100}{15}\left(\frac{1}{6}\right)^{15}\left(\frac{5}{6}\right)^{85} \approx .01002$.
But now let's try to use the previous slide to approximate $P(X = 15)$ via Central Limit Theorem. We note that X is a descrete random variable and normal distribution is continuous one. We need a small "trick" to make them work together:

$$P(X = 15) = P(14.5 < X < 15.5) = P\left(\frac{14.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < \frac{15.5 - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}}\right)$$

$$= P\left(-.58 < \frac{X - 100/6}{\sqrt{100 * \frac{1}{6} * \frac{5}{6}}} < -.31\right) \approx \Phi(-.31) - \Phi(-.58)$$

Where

$$\Phi(t) = P(N(0,1) < t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{y^2}{2}}\,dy.$$

We use tables for $\Phi(t)$ (see Lecture 5.2) we get

$$P(X = 15) \approx 0.1011$$

One more example.

Let X_i, $i = 1, 2, \ldots, 10$ be independent random variables, each being uniformly distributed over $(0, 1)$. Approximate $P\left(\sum_{i=1}^{10} X_i > 7\right)$.

One more example.

Let X_i, $i = 1, 2, \dots, 10$ be independent random variables, each being uniformly distributed over $(0, 1)$. Approximate $P\left(\sum_{i=1}^{10} X_i > 7\right)$.

Solution: We note that $EX_i = .5$ and $Var(X_i) = \frac{1}{12}$, Let X_i, $i = 1, 2, \ldots, 10$ be independent random variables, each being uniformly distributed over $(0, 1)$. Approximate $P\left(\sum_{i=1}^{10} X_i > 7\right)$.

Solution: We note that $EX_i = .5$ and $Var(X_i) = \frac{1}{12}$, thus applying central limit theorem we get

$$P\left(\sum_{i=1}^{10} X_i > 7\right) = P\left(\frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{10 * \frac{1}{12}}} > \frac{7-5}{\sqrt{10 * \frac{1}{12}}}\right)$$

Let X_i, $i = 1, 2, \ldots, 10$ be independent random variables, each being uniformly distributed over $(0,1)$. Approximate $P\left(\sum_{i=1}^{10} X_i > 7\right)$.

Solution: We note that $EX_i = .5$ and $Var(X_i) = \frac{1}{12}$, thus applying central limit theorem we get

$$P\left(\sum_{i=1}^{10} X_i > 7\right) = P\left(\frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{10 * \frac{1}{12}}} > \frac{7-5}{\sqrt{10 * \frac{1}{12}}}\right)$$

$$\approx P(N(0,1) > 2.2)$$

One more example.

Let X_i, $i = 1, 2, \dots, 10$ be independent random variables, each being uniformly distributed over $(0,1)$. Approximate $P\left(\sum_{i=1}^{10} X_i > 7\right)$.

Solution: We note that $EX_i = .5$ and $Var(X_i) = \frac{1}{12}$, thus applying central limit theorem we get

$$P\left(\sum_{i=1}^{10} X_i > 7\right) = P\left(\frac{\sum\limits_{i=1}^{10} X_i - 5}{\sqrt{10 * \frac{1}{12}}} > \frac{7-5}{\sqrt{10 * \frac{1}{12}}}\right)$$

$$\approx P(N(0,1) > 2.2)$$

$$= 1 - P(N(0,1) < 2.2) = 1 - \Phi(2.2) = .0139.$$