It all comes from center, spread, and shape.

1) randomization: this will ensure that the sampling was unbiased. If that is true, the sampling distribution (all possible samples of that size) will have exactly p as its mean. This condition is about CENTER.

2) 10% condition: this will ensure that there is (almost) independence. As long as the population is at least 10 times larger than the sample size, the probability of picking every element of the sample without replacement will not change too much, so the selections are (almost) independent. So that means that we can use the formula for the standard deviation : squareroot(p(1-p)/n). If we did not have that condition, the formula would not apply. So this condition is about SPREAD.

3) *np* and *nq* are both at least 10: Since now we have a formula for standard deviation, we would be able to use the Normal Distribution if our sampling distribution was approximately normal. The sampling distribution for proportions is approximately normal and is closer and closer to a normal distribution when n is large. The rule *np* and *nq* are both at least 10 is a rule of thumb that ensures the SHAPE of the distribution will be close to normal within three standard deviations...

The np > 10 and n(1-p) > 10 come from using the normal approximation to compute a confidence interval. We are relying on the Central Limit Theorem for the approximation and the asymmetry of the binomial distribution suggests that the appropriate sample size n depends on the value of p. If p is close to .5, then smaller samples are needed. This is best seen through simulation. If p is close to 0 or 1, then larger samples are needed.

One explanation for the value of 10 is to say that if we are estimating a proportion, we want the estimates to be in the interval [0, 1]. So, all of the predicted values within 3 standard deviations of our estimate should fit into [0, 1].

So, we need a sample size large enough for $p - 3\sqrt{p(1-p)/n} > 0$ and $p + 3\sqrt{p(1-p)/n} < 1$. Solve for n.

If $p - 3\sqrt{p(1-p)/n} > 0$ then $p > 3\sqrt{p(1-p)/n}$ so $p^2 > 9p(1-p)/n$ and $p > 9(1-p)/n$.

So, np > 9(1-p) will do. But since 1-p < 1 and 10 is such an easy number to remember, if np > 10 then np > 9(1-p) and everything is covered.

Some texts use np > 5. That's just the same analysis using 2 standard deviations instead of 3. The result is np > 4(1-p) so we use np > 5 to be safe and to make it easy to remember.

The upper bound works the same way.