do ∞ var example (13)

<u>Example</u> roll 1 dice. $\mathbb{E}(X_1) = 3.5$ $Var(X_1) = 35/12 \approx 2.9166..$ $sd(X_1) = \sqrt{2.916..}$ 1.707825 $\sqrt{3}$.

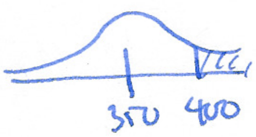
2 dice $\mathbb{E}(X_1+X_2) = 7$ $Var(X_1) = 70/12 \approx 5.8333$ $sd(X_1+X_2) \sim \sqrt{6}$. 2.415229

6 dice $\mathbb{E}(X_1+\dots+X_6) = 21$ $Var(X_1+\dots+X_6) = \frac{35}{2}$ 17.5. $sd(X_1+\dots+X_6) \sim \sqrt{18}$ 4.1833

100 dice $\mathbb{E}(X_1+\dots+X_{100}) = 350$ $var(X_1+\dots+X_{100}) = \frac{3500}{12}$ 291.667 sd $\sqrt{\frac{3500}{12}}$. 17.07825

<u>fact</u> this is well approximated by a normal distribution w/ mean 350 sd $\sqrt{\frac{3500}{12}}$.

<u>Q</u>: estimate appox prob you get ~~between 15 and 20 6's~~ a sum ~~between~~ above 400?

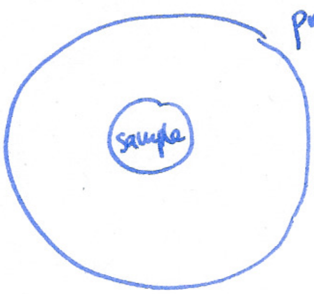


$1 - \text{pnorm}\left(\frac{350-400}{\sqrt{3500/12}}\right)$.

Q: prob get between 10 and 15 6's?

Q: difference between rolling two dice and doubling one die?

§5.2 <u>Sampling distribution of a sample mean</u>



population. ← this has same distribution. (which we may not know)

- if the sample has size 1, it has exactly the distribution of the ~~sample on~~ population, (in particular, same mean and variance)

key facts:
- larger sample size, less variance.
- larger sample size, more normal distribution.

<u>Example</u>. roll 1 dice mean 3.5 var $\sqrt{3}$. s.d. 1.707825

2 dice mean $\frac{7}{2} = 3.5$ var $\sqrt{70/4}$ 1.458333 s.d. 1.207615

6 dice mean $\frac{21}{6} = 3.5$ var $(35/2)/6^2$ s.d. 0.697216

~~100 dice~~ 3.5. ~~0.08101..~~ ~~0.6972167~~ 0.486111

100 dice 3.5 var = $\frac{35}{2}\times 100/100^2$ 0.175 0.41833.

key facts: let $\bar{x}$ be the sample mean of a simple random sample of size n from a population with mean μ, standard deviation σ

then the distribution of $\bar{x}$ has a mean and standard deviation.

(this is not the mean ~~of~~ and standard deviation of the sample!).

- $\mu_{\bar{x}} = \mu$. use $\bar{x}$ as our estimate for μ.
- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. use $\frac{\sigma}{\sqrt{n}}$ to estimate accuracy of estimate.

Example you take an SRS of size 64 from a population w/ mean 37 and standard deviation 16.

$\mu_{\bar{x}} = 37$

$\sigma_{\bar{x}} = \frac{16}{\sqrt{64}} = 2$.

Q suppose you want $\sigma_{\bar{x}} = 1$? How big does n have to be?
↑ size of sample.

We know $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$. what is the distribution of $\mu_{\bar{x}}$?

Central limit theorem

~~Let~~ Take an SRS of size n from a population with mean μ and s.d. $\sigma < \infty$, for large n, $\bar{x}$ is distributed approx $N(\mu, \frac{\sigma}{\sqrt{n}})$.

Special case if population has a normal dist. $\bar{x}$ is exactly normally dist.

Example Suppose heights are dist $N(70,3)$

What is s.d. of sample of size 10? of size 100?

Suppose we want ~~to estimate~~ the sample to be within 1 ~~sd~~ inch of the actual value with 95% probability, how big does n have to be?

Q: what about 0.1 in?

5.3 Sampling distributions for counts and proportions

Yes/no questions : toss coin 100 times, proportion of heads?
will you vote for the incumbent?

Binomial distribution $B(n,p)$ n = # of observations
p = probability of success.

e.g. toss coin 100 times has dist. $B(100, \frac{1}{2})$.

Sampling distribution of counts.

overall population very large ($\gg 20\times$ size of sample) and contains a proportion p of successes. A sample of size n has approx dist $B(n,p)$.

Facts for $X = B(n,p)$ mean $\mu_X = np$

variance $\sigma_X^2 = np(1-p)$

standard deviation $\sigma_X = \sqrt{np(1-p)}$

Warning: counts vs proportions.

X $\qquad \hat{p} = \frac{X}{n}$ mean $\mu_{\hat{p}} = p$.

variance $\sigma_X^2 = \frac{p(1-p)}{n}$

s.d. $\sigma_X = \sqrt{\frac{p(1-p)}{n}}$

R can compute with $B(n,p)$, but if $np \geq 10$ and $n(1-p) \geq 10$

then $X \underset{\text{approx dist}}{\sim} N(np, \sqrt{np(1-p)})$ $\qquad \hat{p} \underset{\text{approx dist}}{\sim} N(p, \sqrt{\frac{p(1-p)}{n}})$.