M358K: October 5th, 2022.

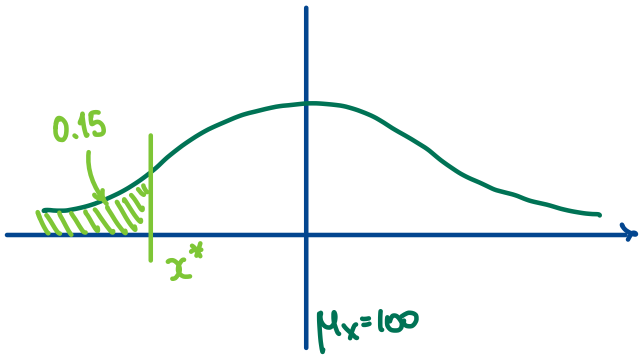
Problem 5.2. (10 points)

At the *Hogwarts School of Witchcraft and Wizardry* the *Ordinary Wizarding Level* (*OWL*) exam is typically taken at the end of the fifth year. Based on hystorical data, we model the *OWL* scores as roughly normal with mean 100 and standard deviation of 16.

$X \sim Normal(mean = 100, sd = 16)$

(a) (5 points)

What is the range of scores for the bottom 15% of the *OWL* takers?



w/ standard normal tables

$z^*_{0.15} = -1.04$

$\downarrow$

$x^* = 100 + 16(-1.04) = 83.36$

or

qnorm(0.15, mean=100, sd=16) = 83.41707

(b) (5 points)
What is the probability that a randomly chosen *OWL* taker has a score higher than 125?

The raw score of 125 corresponds to this score in standard units:

$$z = \frac{125 - 100}{16} = \frac{25}{16} = 1.5625$$

In the std normal tables: $\Phi(1.56) = 0.9406$

$\Rightarrow$ answer: $1 - 0.9406 = 0.0594$

or

1 - pnorm(1.5625) = 0.059

1 - pnorm(125, mean = 100, sd = 16) = 0.059

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 6

The Normal Approximation to the Binomial.

For $Y \sim Binomial(n, p)$ we know that its probability mass function is:

$$p_Y(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, 1, \ldots, n$$

Moreover, its expectation and its variance are

$$\mathbb{E}[Y] = np \quad \text{and} \quad Var[Y] = np(1-p).$$

Now, consider a sequence of binomial random variables $Y_n \sim Binomial(n, \boxed{p})$. Note that, while the number of trials n varies, the probability of success in every trial p remains the same for all n. The *normal approximation to the binomial* is a theorem which states that

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \stackrel{\mathcal{D}}{\Rightarrow} N(0, 1)$$

Practically, this means that Y_n is "approximately" normal with mean np and variance $np(1-p)$ for "large" n. The usual rule of thumb is that both $np > 10$ and $n(1-p) > 10$.

Another practical adjustment needs to be made due to the fact that discrete distributions of Y_n are approximated by a continuous (normal) distribution. This adjustment is usually referred to as the **continuity correction**. More specifically, provided that the conditions above are satisfied, for every integer $a < b$, we have that

$$\mathbb{P}[a \leq Y_n \leq b] = \mathbb{P}\left[a - \tfrac{1}{2} < Y_n < b + \tfrac{1}{2}\right]$$

$$= \mathbb{P}\left[\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}} < \boxed{\frac{Y_n - np}{\sqrt{np(1-p)}}} < \frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right] \approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}}\right)$$

$\approx N(0,1)$

where Φ, as usual, stands for the cummulative distribution function of the standard normal distribution.

For more about the history of the theorem and ideas for its proof, go to: Wikipedia: de Moivre-Laplace.

Problem 6.1. A student takes an exam with 200 TRUE/FALSE questions. Shirley knows the correct answer to exactly 100 questions. For the remaining questions, she guesses at random. The passing mark is $\boxed{136}$ correct answers. What is the (approximate) probability she passes the exam?

Y... # of correct guesses

$Y \sim Binomial(n = 100, p = 0.5)$ $\quad n \cdot p = n(1-p) = 50 > 10$ ✓

We seek: $\mathbb{P}[Y \geq 36] = ?$ $\quad \mathbb{E}[Y] = 50 \quad SD[Y] = \sqrt{100(0.5)(0.5)} = 5$

$$\mathbb{P}[Y \geq 36] = 1 - \mathbb{P}[Y \leq 35] = 1 - \Phi\left(\frac{35.5 - 50}{5}\right)$$

$$= 1 - \Phi(-2.9)$$

standard normal tables (SNT) : $1 - 0.0019 = 0.9981$

1 - pnorm(-2.9) = 0.9981342

1-pnorm(35.5, mean=50, sd=5) = 0.9981342

or 1-pbinom(35, 100, 0.5) = 0.9982412