Note # 5: Probability Distributions

Problem 1. Heights of adult males ages 20-24 are approximately normally distributed with a mean of 70 inches and a standard deviation of 5. Using this and the table of z-scores:

- **a.** What percent of men in this age range are shorter than 67 inches?
- **b.** What percent of men in this age range are taller than 54 inches?
- **c.** What percent of men in this age range are shorter than 76.5 inches?
- **d.** What percent of men in this age range are taller than 72 inches?
- **e.** What percent of men in this age range are between 68 and 73 inches tall?
- **f.** What percent of men in this age range are between 63 and 67 inches tall?
- **g.** What percent of men in this age range are between 74 and 86 inches tall?
- **h.** What height is greater than that of 75% of all adult males in this age range?
- **i.** What height is less than that of 45% of all adult males in this age range?
- **j.** Between what two heights do the middle 95% of all adult males in this age range heights fall?
- **k.** Your friend tells you that he is shorter than 5% of adult males in this age range. How tall is he?

Solution:

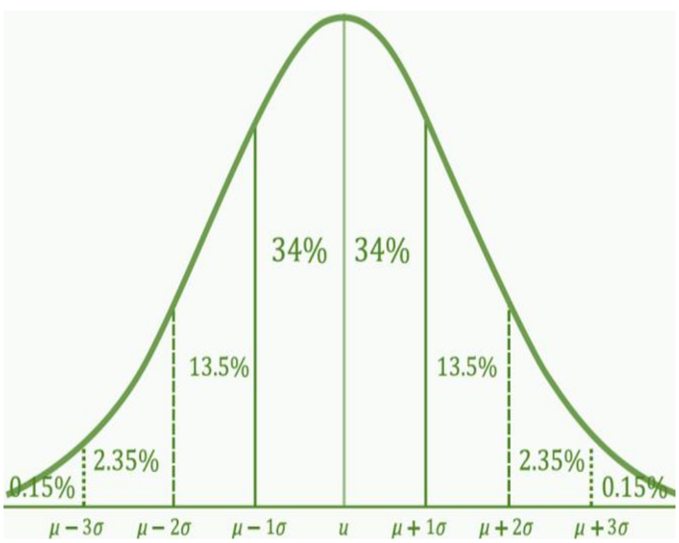
- **a.** $P(X < 67) = P\left(Z < \frac{67-70}{5}\right) = P\left(Z < \frac{-3}{5}\right) = P(Z < -0.60) = 0.2743 = \mathbf{27.43}\%$
- **b.** $P(X > 54) = P\left(Z > \frac{54-70}{5}\right) = P\left(Z > \frac{-16}{5}\right) = P(Z > -3.20) = 1 - P(Z < -3.20) = 1 - 0.0007 = 0.9993 = \mathbf{99.93}\%$
- **c.** $P(X < 76.5) = P\left(Z < \frac{76.5-70}{5}\right) = P\left(Z < \frac{6.5}{5}\right) = P(Z < 1.30) = 0.9032 = \mathbf{90.32}\%$
- **d.** $P(X > 72) = P\left(Z > \frac{72-70}{5}\right) = P\left(Z > \frac{2}{5}\right) = P(Z > 0.40) = 1 - P(Z < 0.40) = 1 - 0.6554 = 0.3446 = \mathbf{34.46}\%$
- **e.** $P(68 < X < 73) = P\left(\frac{68-70}{5} < Z < \frac{73-70}{5}\right) = P\left(\frac{-2}{5} < Z < \frac{3}{5}\right) = P(-0.40 < Z < 0.60) = P(Z < 0.60) - P(Z < -0.40) = 0.7257 - 0.3446 = 0.3811 = \mathbf{38.11}\%$
- **f.** $P(63 < X < 67) = P\left(\frac{63-70}{5} < Z < \frac{67-70}{5}\right) = P\left(\frac{-7}{5} < Z < \frac{-3}{5}\right) = P(-1.40 < Z < -0.60) = P(Z < -0.60) - P(Z < -1.40) = 0.2743 - 0.0808 = 0.1935 = \mathbf{19.35}\%$
- **g.** $P(74 < X < 86) = P\left(\frac{74-70}{5} < Z < \frac{86-70}{5}\right) = P\left(\frac{4}{5} < Z < \frac{16}{5}\right) = P(0.80 < Z < 3.20) = P(Z < 3.20) - P(Z < 0.80) = 0.9993 - 0.7881 = 2112 = \mathbf{21.12}\%$
- **h.** $P(X < ??) = 0.75 \;\rightarrow 0.67 = \frac{obs-70}{5} = obs = (0.67)(5) + 70 = \mathbf{73.35\ in}$
- **i.** $P(X < ??) = 0.55 \;\rightarrow 0.13 = \frac{obs-70}{5} = obs = (0.13)(5) + 70 = \mathbf{70.65\ in}$
- **j.** $X_1: P(X < ??) = 0.025 \;\rightarrow -1.96 = \frac{obs-70}{5} = obs = (-1.96)(5) + 70 = \mathbf{60.2\ in}$ and
  $X_2: P(X < ??) = 0.975 \;\rightarrow 1.96 = \frac{obs-70}{5} = obs = (1.96)(5) + 70 = \mathbf{79.8\ in}$
  Between 60.2 and 79.8 inches
- **k.** $P(X < ??) = 0.95 \;\rightarrow 1.645 = \frac{obs-70}{5} = obs = (1.645)(5) + 70 = \mathbf{78.225\ in}$. We wouldn't be able to get this value from the table, you would need to use technology (not requested for this course).

Problem 2. The heights of children ages 3-5 are approximately normally distributed with a mean of 40 inches and a standard deviation of 2.5 inches. Use this information and the Empirical rule to answer the following questions:

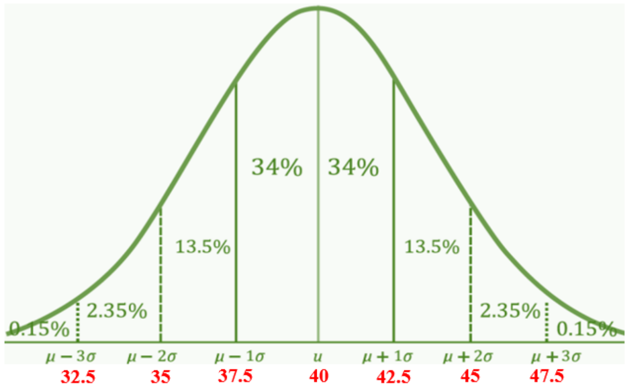
- **a.** Draw the appropriate Empirical Rule diagram for this scenario.
- **b.** What percent of children are between 40 and 45 inches tall?
- **c.** What percent of children are above 42.5 inches tall?
- **d.** What percent of children are between 32.5 inches tall and 42.5 inches tall?
- **e.** What percent of children are below 35 inches tall?
- **f.** What percent of children are below 40 inches tall?
- **g.** What percent of children are above 37.5 inches tall?
- **h.** What percent of children are either less than 35 inches tall or more than 42.5 inches tall?

Solution:

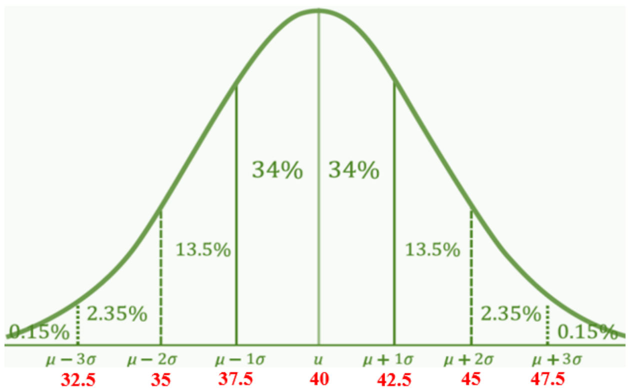
- **a.** Empirical Rule diagram:



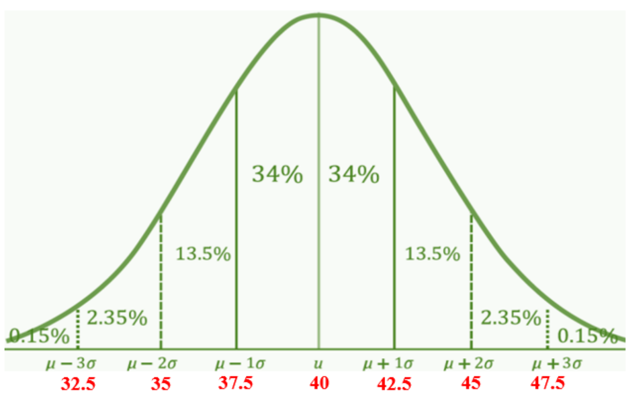
b. $\%$ between 40 & $45 = 34\% + 13.5\% = \mathbf{47.5}\%$



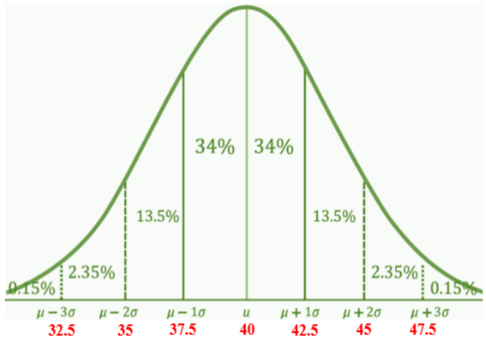
c. $\%$ above $42.5 = 13.5\% + 2.35\% + 0.15\% = \mathbf{16}\%$



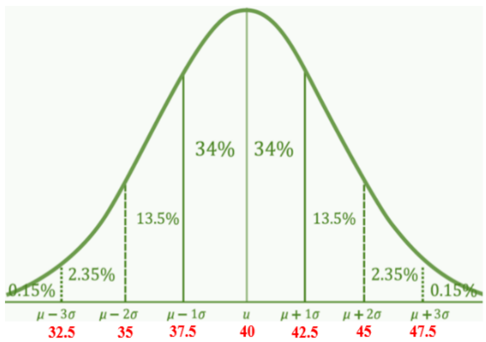
d. $\%$ between 32.5 & $42.5 = 2.35\% + 13.5\% + 34\% + 34\% = \mathbf{83.85}\%$



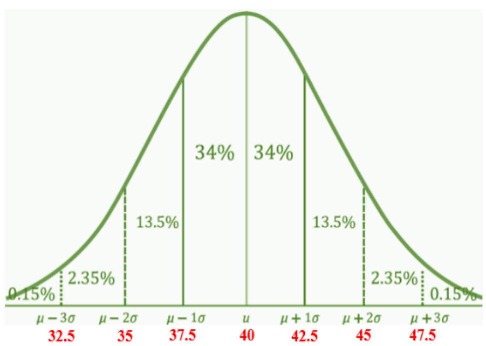
e. % below 35 $= 0.15\% + 2.35\% += \mathbf{2.5}\%$



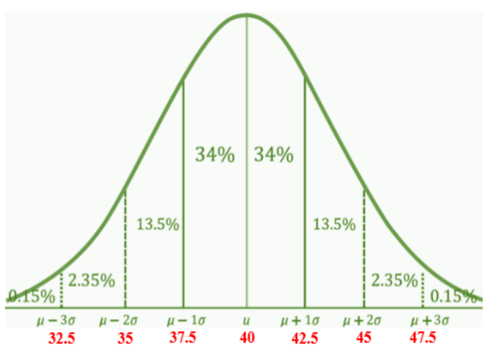
f. % above 40 $= 0.15\% + 2.35\% + 13.5\% + 34\% = \mathbf{50}\%$



g. % above 37.5 $= 34\% + 34\% + 13.5\% + 2.35\% + 0.15\% = \mathbf{84}\%$



h. Less than 35 or greater than 42.5 $= 0.15\% + 2.35\% + 13.5\% + 2.35\% + 0.15\% = 18.5\%$



Problem 3. The Wechsler Adult Intelligence Scale (WAIS) is an IQ test. Scores on the WAIS for the 20 to 34 age group are approximately Normally distributed with a mean of 110 and a standard deviation of 15. Scores for the 60 to 64 age group are approximately Normally distributed with a mean of 90 and a standard deviation of 15. Sarah, who is 30, scores 130 on the WAIS. Her mother, who is 60, takes the test and scores 110. Express both scores as standard scores that show where each woman stands within her own age group. Who scored higher relative to her age group, Sarah or her mother?

Solution:

Sarah (30)	**Mother (60)**
$X \sim N(\mu = 110, \sigma = 15)$ $obs = 130$ $Z = \frac{obs - mean}{SD} = \frac{130 - 110}{15} = \frac{20}{15} = 1.33$	$X \sim N(\mu = 90, \sigma = 15)$ $obs = 110$ $= \frac{obs - mean}{SD} = \frac{110 - 90}{15} = \frac{20}{15} = 1.33$

Standardized scores are the same → relative to their individual age group, they performed the same.

Problem 4. SAT scores (out of 2400) are distributed normally with a mean of 1500 and a standard deviation of 300. Suppose a school council awards a certificate of excellence to all students who score at least 1900 on the SAT. We randomly pick one of the recognized students.

- **a.** What proportion of students are recognized?
- **b.** What is the probability that the randomly selected student scored at least 1900?
- **c.** What is the probability that the randomly selected student scored at least 2150?
- **d.** What is the probability that the randomly selected student scored between 2000 and 2100?
- **e.** What is the probability that the randomly selected student scored between 1900 and 2100?

Solution:

a. $P(recognized) = P(X > 1900)$

$$Z = \frac{obs - mean}{SD} = \frac{1900 - 1500}{300} = \frac{400}{300} = 1.33$$

$$P(X > 900) = P(Z > 1.33) = 1 - P(Z < 1.33) = 1 - 0.9082 = \mathbf{0.0918}$$

b. $P(randomly\ selected\ student\ scored\ at\ least\ 1900)$

$$P(\ scored\ at\ least\ 1900|recognized) = P(\ scored\ at\ least\ 1900|scored\ at\ least\ 1900) = \mathbf{1}$$

c. $P(\ scored\ at\ least\ 2150|scored\ at\ least\ 1900) = \frac{P(scored\ at\ least\ 2150\ and\ \ scored \geq 1900)}{P(scored \geq 1900)}$

$$= \frac{P(scored\ at\ least\ 2150\ )}{P(scored \geq 1900)}$$

$$Z = \frac{2150-1500}{300} = \frac{650}{300} = 2.17$$

$$\frac{P(X \geq 2150)}{P(X \geq 1900)} = \frac{0.015}{0.0918} = \mathbf{0.1634}$$

d. $P(scored\ between\ 2000\ and\ 2100|scored\ \geq 1900) = \frac{P(2000<X<2100)}{P(scored \geq 1900)}$

$$= \frac{P\left(\frac{2000-1500}{300}<Z<\frac{2100-1500}{300}\right)}{P\left(Z>\frac{1900-1500}{300}\right)} = \frac{P(1.67<Z<2.00)}{P(Z>1.33)} = \frac{0.9772-0.9525}{0.0918} = \mathbf{0.2691}$$

e. $P(scored\ between\ 1900\ and\ 2100|scored\ \geq 1900) = \frac{P(1900<X<2100)}{P(scored \geq 1900)}$

$$= \frac{P\left(\frac{1900-1500}{300}<Z<\frac{2100-1500}{300}\right)}{P\left(Z>\frac{1900-1500}{300}\right)} = \frac{P(1.33<Z<2.00)}{P(Z>1.33)} = \frac{0.9772-0.9082}{0.0918} = \mathbf{0.7516}$$

Problem 5. A professional basketball player has a 73% success rate when shooting free throws. Let X represent the number of free throws he makes in a random sample of 5 free throws. Assume that these free throws fit the requirements for a binomial experiment.

- **a.** What is the probability that he makes exactly 1 free throw? Write out the formula that allows you to get to this answer.
- **b.** What is the probability that he will make less than 2 free throws?
- **c.** What is the probability that he will make between 1 and 3 free throws, inclusive?
- **d.** What is the expected number of free throws that he will make?
- **e.** What is the variance of X?
- **f.** What is the standard deviation of X?

Solution:

$$n = 5,\ \ p = 0.73, 1 - p = 1 - 0.73 = 0.27,\ \ \ k = 1$$

a. $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \ = \ P(X = 1) = \binom{5}{1} (0.73)^1 (1 - 0.73)^{5-1}$

$= \frac{5!}{1! \times 4!}(0.73)(0.27)^4 = \mathbf{0.019}$

b. $P(X < 2) = P(X = 0) + P(X = 1)$

$\because P(X = 0) = \binom{5}{0} (0.73)^0 (1 - 0.73)^{5-0} = 0.0014$, from part a, we know $P(X = 1) = 0.019$

$\therefore P(X < 2) = 0.0014 + 0.019 = \mathbf{0.0204}$

c. $P(1 \le X \le 3) = P(X = 1) + P(X = 2) + P(X = 3)$

$P(X = 1) = 0.019,$

$P(X = 2) = \binom{5}{2} (0.73)^2 (1 - 0.73)^3 = 0.105,$

$P(X = 3) = \binom{5}{3} (0.73)^3 (1 - 0.73)^2 = 0.284$

$\therefore P(1 \le X \le 3) = 0.019 + 0.105 + 0.284 = \mathbf{0.408}$

d. $E(X) = \mu = n \times p = 5 \times 0.73 = \mathbf{3.65}$

e. $var(X) = n \times p \times (1 - p) = 5 \times 0.73 \times 0.27 = \mathbf{0.9855}$

f. $Stdev = \sqrt{n \times p \times (1 - p)} = \sqrt{5 \times 0.73 \times 0.27} = \sqrt{0.9855} = \mathbf{0.9927}$

Problem 6. While it is often assumed that the probabilities of having a boy or a girl are the same, the actual probability of having a boy is slightly higher at 0.51. Suppose a couple plans to have 3 kids.

a. Use the binomial model to calculate the probability that two of them will be boys.

b. Write out all possible orderings of 3 children, 2 of whom are boys. Use these scenarios to calculate the same probability from part (a) but using the addition rule for disjoint outcomes. Confirm that your answers from parts (a) and (b) match.

c. If we wanted to calculate the probability that a couple who plans to have 8 kids will have 3 boys, briefly describe why the approach from part (b) would be more tedious than the approach from part (a).

Solution:

a. Binomial Distribution, $n = 3, \quad p = 0.51, \quad 1 - p = 0.49$

$$P(X = 2) = \binom{3}{2} (0.51)^2 \, (1 - 0.49)^1 = \mathbf{0.3823}$$

b. $P(B, B, G) = 0.51 \times 0.51 \times 0.49 = \mathbf{0.12744}$

$P(B, G, B) = 0.51 \times 0.49 \times 0.51 = \mathbf{0.12744}$

$P(G, B, B) = 0.49 \times 0.51 \times 0.51 = \mathbf{0.12744}$

$P(2\ Boys) = 0.12744 + 0.12744 = \mathbf{0.3823}$

c. $\binom{8}{3} = 56$, very tedious to write out all 56 scenarios, may miss some.

Problem 7. Suppose a university announced that it admitted 2,500 students for the following year's freshman class. However, the university has dorm room spots for only 1,786 freshman students. If there is a 70% chance that an admitted student will decide to accept the offer and attend this university, what is the approximate probability that the university will not have enough dormitory room spots for the freshman class?

Solution:

$X = \#enroll \to X \sim Binomial\ (n = 2500, p = 0.70)$

$P(X \geq 1787) = P(X = 1787) = P(X = 1788) + \cdots + P(X = 2500)$

Normal distribution to approximate Binomial distribution:

$n \times p = 2500 \times 0.70 = \mathbf{1750} \geq \mathbf{10}$

$n \times (1 - p) = 2500 \times 0.30 = \mathbf{750} \geq \mathbf{10}$

$Mean = \mu = E(X) = n \times p = 2500 \times 0.70 = \mathbf{1750}$

$Stdev = \sqrt{n \times p \times (1 - p)} = \sqrt{2500 \times 0.70 \times 0.30} = \mathbf{23}$

$$Z = \frac{obs - mean}{SD} = \frac{1787 - 1750}{23} = \frac{37}{23} = 1.61$$

$P(Z > 1.61) = 1 - P(Z < 1.61) = 1 - 0.9463 = \mathbf{0.0537}$

Problem 8. Pew Research reported that the typical response rate to their surveys is only 9%. If for a particular survey 15,000 households are contacted, what is the probability that at least 1,500 will agree to respond?

Solution:

$X = \#who\ respond \rightarrow X \sim Binomial\ (n = 15000, p = 0.09)$

$P(X \geq 1500) = ?$

Normal distribution to approximate Binomial distribution:

$n \times p = 15000 \times 0.09 = \mathbf{1350 \geq 10}$

$n \times (1 - p) = 15000 \times 0.91 = \mathbf{13650 \geq 10}$

$Mean = \mu = E(X) = n \times p = 15000 \times 0.09 = \mathbf{1350}$

$Stdev = \sqrt{n \times p \times (1 - p)} = \sqrt{15000 \times 0.09 \times 0.91} \approx \mathbf{35}$

$$Z = \frac{obs - mean}{SD} = \frac{1500 - 1350}{35} = \frac{150}{35} = 4.29$$

$P(Z > 1500) = 1 - P(Z < 4.29) = 1 - approx\ 1 \approx \mathbf{0}$