

Review Problems for Exam 3

Math 1040–1

1. Find the probability that a standard normal random variable is less than 2.37.

Looking up 2.37 on the normal table, we see that the probability is 0.9911.

2. Find the probability that a standard normal random variable is at least -1.04 .

Looking up -1.04 on the table, we see that the probability is $1 - 0.1492 = 0.8508$.

3. Find the probability that a standard normal random variable is between 0.56 and 2.83.

Looking up 2.83 and 0.56, we see that the probability is $0.9977 - 0.7123 = 0.2854$.

4. Find the probability that a standard normal random variable is exactly 0.01.

The probability of a normal random variable taking exactly one value is always zero.

5. Porphyrin is a pigment in blood protoplasm and other body fluids that is significant in body energy and storage. Let x be a random variable that represents the number of milligrams of porphyrin per deciliter of blood. In healthy adults, x is approximately normally distributed with mean 38 and standard deviation 12.

- (a) What is the probability that x is less than 50?

First we will calculate the z -score: $z = \frac{50 - 38}{12} = 1$. Then, we look up 1 in the table to see that the probability is 0.8413.

- (b) What is the probability that x is between 16 and 60?

The z -scores for 16 and 60 are:

$$z = \frac{16 - 38}{12} = -1.83 \qquad z = \frac{60 - 38}{12} = 1.83$$

Looking up these probabilities in the table, we see that the probability of being between the two numbers is $0.9664 - 0.0336 = 0.9328$.

- (c) What is the probability that x is more than 60?

The z -score for 60 is 1.83 (as we calculated in the last part), so the probability of being above this number is $1 - 0.9664 = 0.0336$.

- (d) What is the 70th percentile for the number of milligrams of porphyrin per deciliter of blood?

The probability closest to 0.7 on the table is 0.6985, which happens at $z = 0.52$. Solving the equation $0.52 = \frac{x - 38}{12}$, we get $x = 44.24$, so the 70th percentile is 44.24.

- (e) What is the minimum amount of porphyrin that will be in the top 10% of observations?

Since 10% of observations are above the number we are looking for, 90% are below. The probability closest to 0.9 on the table is 0.8997, which occurs at $z = 1.28$. Solving the equation $1.28 = \frac{x - 38}{12}$, we get $x = 53.36$.

6. Coal is carried from a mine in West Virginia to a power plant in New York in hopper cars on a long train. Each car has a mean weight of 75 tons and a standard deviation of 0.8 tons. The weights of the cars are approximately normal.

- (a) What is the probability that a car chosen at random will have less than 74.75 tons of coal?

The z -score for 74.75 is $z = \frac{74.75 - 75}{0.8} = -0.31$, so the probability of being less than this number is 0.3783.

- (b) What is the probability that 40 cars chosen at random will have a mean weight of less than 74.75 tons of coal?

The standard deviation of the sample mean is $\sigma_{\bar{x}} = 0.8/\sqrt{40} = 0.12649$, so the z -score is $z = \frac{74.75 - 75}{0.12649} = -1.98$. The probability from the table is 0.0239.

- (c) What is the probability that 40 cars chosen at random will have a mean weight of at least 75.8 tons of coal?

Again, we are asked about the mean, so the z -score is $z = \frac{75.8 - 75}{0.12649} = 6.32$. Since this number is too large to be on the table, we can assume that the number the table would give is 1. Therefore, the probability of getting at least this number is $1 - 1 = 0$.

7. The *Denver Post* stated that 80% of all new products introduced in grocery stores are taken off the market within 2 years. If a grocery chain introduces 66 new products, what is the probability that within 2 years

(a) 47 or more fail?

Since $np \geq 5$ and $nq \geq 5$, we can use the normal approximation for the binomial distribution. The mean is $np = 66 \cdot 0.8 = 52.8$ and the standard deviation is $\sqrt{66 \cdot 0.8 \cdot 0.2} = 3.2496$. Using the continuity correction, we will be looking at the probability of a normal random variable being at least 46.5. The z -score is $z = \frac{46.5 - 52.8}{3.2496} = -1.94$, so the probability is $1 - 0.0262 = 0.9738$.

(b) 58 or fewer fail?

With the continuity correction, we will consider 58.5. The z -score is $z = \frac{58.5 - 52.8}{3.2496} = 1.75$, so the probability is 0.9599.

(c) 15 or more succeed?

“15 or more succeeding” is the same event as “51 or fewer failing.” The continuity correction makes us consider a normal random variable being less than 51.5, so the z -score is $z = \frac{51.5 - 52.8}{3.2496} = -.40$ and the probability is 0.3446.

(d) fewer than 10 succeed?

“Fewer than 10 succeeding” is the same event as “more than 56 failing,” which is the same as “at least 57 failing.” We will need to consider a normal random variable being above 56.5, so the z -score is $z = \frac{56.5 - 52.8}{3.2496} = 1.14$. the probability of being above this value is $1 - 0.8729 = 0.1271$.

8. The number of days each year that the Salt Lake City National Weather Service Office's forecast has been off by at least 10 degrees Fahrenheit are:

Year	1950	1960	1970	1980	1990	2000	2010
Number of Days	55	41	23	13	8	9	3

- (a) Draw a scatter plot for the above data.

On next page

- (b) Is the linear correlation positive, negative, or neither? Explain.

The correlation is negative, since the y -values tend to decrease as x increases.

- (c) Find the correlation coefficient, r .

We will need to use the formula $r = \frac{n \sum(xy) - (\sum x)(\sum y)}{\sqrt{n \sum(x^2) - (\sum x)^2} \sqrt{n \sum(y^2) - (\sum y)^2}}$.

Let x represent the number of decades since 1950 (so, $x = 0, 1, 2, \dots, 6$). Then, $\sum x = 21$, $\sum y = 152$, $\sum(xy) = 221$, $\sum(x^2) = 91$, $\sum(y^2) = 5558$, and $n = 7$. Then,

$$r = \frac{7 \cdot 221 - 21 \cdot 152}{\sqrt{7 \cdot 91 - 21^2} \sqrt{7 \cdot 5558 - 152^2}} = -0.935.$$

- (d) Write the equation of the regression line and sketch the line on the scatter plot from part 8a.

To find the slope of the regression line, we need the formula

$$m = \frac{n \sum(xy) - (\sum x)(\sum y)}{n \sum(x^2) - (\sum x)^2} = \frac{7 \cdot 221 - 21 \cdot 152}{7 \cdot 91 - 21^2} = -8.393.$$

The y -intercept is given by the formula

$$b = \frac{\sum y}{n} - m \frac{\sum x}{n} = \frac{152}{7} - (-8.393) \frac{21}{7} = 46.893.$$

Therefore, the equation of the regression line is $\hat{y} = -8.393x + 46.893$, where x represents the number of decades since 1950.

- (e) Predict the number of days the SLC NWS's forecast will be off by at least 10 degrees in 2020.

Since x is given in decades since 1950, 2020 is represented by $x = 7$. Using our equation from the last part, $\hat{y} = -8.393 \cdot 7 + 46.893 = -11.858$ is the prediction for 2020 (Note: The answer does not make sense in the context of the question since we are measuring something that cannot be negative. The negative result is partly due to extrapolation and partly due to the data looking closer to exponential than linear in the scatter plot.).

- (f) Calculate the residual for the 2010 observation.

The residual, d , is calculated by $d = y - \hat{y}$. Using $x = 6$, we get that $\hat{y} = -8.393 \cdot 6 + 46.893 = -3.465$, so the residual is $d = 3 - (-3.465) = 6.465$.

Graph for 8a and 8d:

