MATH 3339 Statistics for the Sciences

Final Review

Wendy Wang
wwang60@central.uh.edu

MC 15 $\times 3 = 45$ 120 min

FR 5 $\times 11 = 55$

Final Review - 3339

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 1

$= P(A \text{ and } B)$
$= P(A \cap B)$

Given that $P(A) = 0.2$, $P(B) = 0.3$, and $P(A\&B) = 0.1$, find $P(A|B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.3} = 0.33$$

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.2 + 0.3 - 0.1 = \boxed{0.4}$$

$P(A) = 0.2$, $P(B) = 0.3$, $P(A \cup B) = 0.4$, find $P(A|B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 2

X: # of ppl. within 1 hr
X ~ poisson(λ = 5/hr)

The number of people arriving for treatment at an emergency room can be modeled by a Poisson process with a mean of five people per hour.

1. What is the probability that exactly four arrivals occur at a particular hour?

   $P(X=4) = \text{dpois}(4, 5)$

2. How many people do you expect to arrive during a 45-min period?

   $\frac{5}{60} \times 45 = 3.75$

3. What is the probability that less than 3 people arrive during a 45-min period?

   Y = # of people during 45 min

   Y ~ poisson(λ = 3.75)

   $P(Y<3) = P(Y=0,1,2)$

   ppois(2, 3.75)

Example 3

In testing a certain kind of missile, target accuracy is measured by the average distance X (from the target) at which the missile explodes. The distance X is measured in miles and the distribution of X is given by:

X	0	10	50	100
P(X)	$\frac{1}{14}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{2}$

$\Sigma P(y) = 1$

Find the mean and variance for the target accuracy.

mean

$\mu = E(x) = \Sigma x P(X = x)$

$= 0(\frac{1}{14}) + 10(\frac{1}{7}) + 50(\frac{2}{7}) + 100(\frac{1}{2})$

$Var(x) = E(x^2) - [E(x)]^2$

$= 0^2(\frac{1}{14}) + 10^2(\frac{1}{7}) + 50^2(\frac{2}{7}) + 100^2(\frac{1}{2}) - [E(x)]^2$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

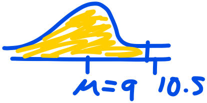
Example 4

$X \sim N(\mu=9, \sigma=3.2)$

The weights of individual bolts produced at a manufacturing plant, X, is normally distributed. If the mean weight of the bolts is 9 grams and the standard deviation is 3.2 grams, find:

1. $P(X \leq 10.5)$ $= P(X<10.5)$

   pnorm(10.5, 9, 3.2)



2. $P(X \geq 7.1)$

   1-pnorm(7.1, 9, 3.2)

3. The value of x such that $P(X \leq x) = 0.93$

   x_0 = qnorm(0.93, 9, 3.2)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 5

$p = 0.05$

In testing a new drug, researchers found that 5% of all patients using it will have a mild side effect. A random sample of 7 patients using the drug is selected. Find the probability that: X = # of patients having side effect in sample of 7.

$X: 1, 2, 3, 4, 5, 6, 7$

$X \sim binom(n=7, p=0.05)$

1. None will have this mild side effect.

$P(X=0) = dbinom(\underset{x}{0}, \underset{n}{7}, \underset{p}{0.05})$

2. Exactly 2 patients will have this mild side effect.

$P(X=2) = dbinom(2, 7, 0.05)$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 5

In testing a new drug, researchers found that 5% of all patients using it will have a mild side effect. A random sample of 7 patients using the drug is selected. Find the probability that:

3. At least one will have this mild side effect.

$P(x \text{ at least } 1) = P(x \text{ can be } 1,2,3,4,5,6,7)$

$= 1 - P(x=0)$

$= 1 - \text{dbinom}(0,7,0.05)$

4. What is the expected value and variance of the number of patients that will have this mild side effect?

x	0	1	2	3	4	5	6	7
$P(X=x)$								?

?

$\mu = n \cdot p = 7 \times 0.05$

$Var(x) = np(1-p) = 7 \times 0.05 \times 0.95$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 6 $Z \sim N(0,1)$

Let Z be the standard normal random variable. Calculate the following.

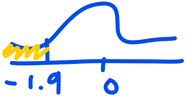
1. $P(|Z| \leq 2.4)$ $= P(-2.4 \leq Z \leq 2.4) = P(Z \leq 2.4) - P(Z < -2.4)$

$=$ pnorm(2.4) $-$ pnorm(-2.4)

or

pnorm(2.4, 0, 1) $-$ pnorm(-2.4, 0, 1)

2. $P(Z \leq -1.9)$



pnorm(-1.9, 0, 1)

3. Find c such that $P(Z \geq c) = 0.98$ $\quad P(Z < c) = 1 - 0.98 = 0.02$

$c =$ qnorm(0.02, 0, 1)

UNIVERSITY of HOUSTON DEPARTMENT OF MATHEMATICS

Example 7

It has been estimated that as many as 70% of the fish caught in certain areas of the Great Lakes have liver cancer due to the pollutants present. Find an approximate 95% range for the percentage of fish with liver cancer present in a sample of 130 fish.

95% CI for p. $\hat{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ ← 150

0.7 + c(-1,1) * qnorm($\frac{1+0.95}{2}$)

Example 8

In a hypothesis test, if the computed P-value is less than 0.001, there is very strong evidence to

$P < \alpha$

reject

a) retest with a different sample.
b) accept the null hypothesis
c) fail to reject the null hypothesis.
d) reject the null hypothesis.

Example 9

1. A simple random sample of 100 8th graders at a large suburban middle school indicated that 86% of them are involved with some type of after school activity. Find the 98% confidence interval that estimates the proportion of them that are involved in an after school activity. $\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

2. An SRS of 24 students at UH gave an average height of 6.1 feet and a standard deviation of .3 feet. Construct a 90% confidence interval for the mean height of students at UH.

   σ is given? NO $s=0.3$ $\bar{x} \pm qt(\frac{\alpha}{2}, df) \cdot \frac{s}{\sqrt{n}}$

3. The average height of students at UH from an SRS of 17 students gave a standard deviation of 2.9 feet. Construct a 95% confidence interval for the standard deviation of the height of students at UH. Assume normality for the data.

   $\left(\frac{(n-1)s^2}{\chi^2} \right. , \sqrt{}$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 10

m: # of vehicles

Data for gas mileage (in mpg) for different vehicles was entered into a software package and part of the ANOVA table is shown below:

model

Source	DF	SS	MS
Vehicle	2	440	220.00
Error	17	318	18.71
Total	19	758	

$\frac{SS}{DF}$ "F" Pr

$19 = 2 + 17 = n-1$

1. Determine the value of the test statistic F to complete the table.

$f = \frac{MS(model)}{MSE} = \frac{220}{18.71} = 11.76$

2. Determine the p-value. $F \sim F(2,17)$

P-value $= P(F > 11.76)$

$1 - P(F < 11.76) = 1 - pf(11.76, 2, 17)$

$= 0.00062 \rightarrow R\ H_0$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 11

left tailed

- The one-sample t statistic for a test of $H_0 : \mu = 12$ vs. $H_a : \mu < 12$ based on $n = 174$ observations has the test statistic value of $t = -1.58$. What is the p-value for this test?

Pvalue

t=1.58

+ curve

$$\text{p-value} = P(t \le -1.58)$$
$$= pt(-1.58, \underbrace{173}_{n-1})$$

- The one-sample t statistic for a test of $H_0 : \mu = 12$ vs. $H_a : \mu > 12$ based on $n = 174$ observations has the test statistic value of $t = 1.58$. What is the p-value for this test?

0 1.58

$$\text{pvalue} = P(t > 1.58)$$
$$= 1 - pt(1.58, 173)$$

Example 12

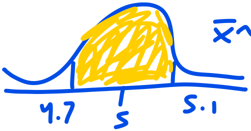
The average life of a manufacturer's blender is 5 years, with a standard deviation of 1 year. Assuming that the lives of these blenders approximately follow a normal distribution, find the probability that the mean life of a random sample of 25 such blenders falls between 4.7 and 5.1 years. $X \sim N(\mu = 5, \sigma = 1)$

$P(4.7 < \bar{X} < 5.1)$ $(n = 25)$

$\bar{X} \overset{CLT}{\sim} N\left(\mu_{\bar{x}} = \mu = 5, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{25}} = 0.2\right)$

$= \text{pnorm}(5.1, 5, 0.2) - \text{pnorm}(4.7, 5, 0.2)$

$\bar{x} \sim N(5, 0.2)$



Example 13

Suppose $f(x, y) = \frac{x+2y}{18},\ x = 1, 2;\ y = 1, 2$ is the joint pmf of X and Y. Determine $P(X + Y = 3)$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 14

Let X be the amount of time (in hours) the wait is to get a table at a restaurant. Suppose the cdf is represented by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4}x^2 & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

1. Find $P(X \leq 1.5)$
2. Find $P(1 \leq X \leq 1.5)$
3. Find the density function $f(x)$.
4. Find the mean and variance of X.

Example 15

Q35 from test review.

Example 16

Below is the computer output for the appraised value (in thousands of dollars) and number of rooms for 20 houses in East Meadow, New York.

Predictor	Coef	Stdev	t-ratio
Constant	74.80	19.04	3.93
Rooms	19.718	2.631	7.49

$S = 29.05$ R-sq = 43.8% R-sq (adj) = 43.0%

1. What is the regression equation?

UNIVERSITY of HOUSTON DEPARTMENT OF MATHEMATICS

Example 16

Below is the computer output for the appraised value (in thousands of dollars) and number of rooms for 20 houses in East Meadow, New York.

Predictor	Coef	Stdev	t-ratio
Constant	74.80	19.04	3.93
Rooms	19.718	2.631	7.49

$S = 29.05$ R-sq = 43.8% R-sq (adj) = 43.0%

2. Predict the price of a 10 room house (in thousands of dollars).

Example 16

Below is the computer output for the appraised value (in thousands of dollars) and number of rooms for 20 houses in East Meadow, New York.

Predictor	Coef	Stdev	t-ratio
Constant	74.80	19.04	3.93
Rooms	19.718	2.631	7.49

$S = 29.05$ R-sq = 43.8%   R-sq (adj) = 43.0%

3. Calculate the 95% confidence interval of the slope of the regression line for all homes.

Example 16

Below is the computer output for the appraised value (in thousands of dollars) and number of rooms for 20 houses in East Meadow, New York.

Predictor	Coef	Stdev	t-ratio
Constant	74.80	19.04	3.93
Rooms	19.718	2.631	7.49

$S = 29.05$ R-sq = 43.8% R-sq (adj) = 43.0%

4. Use the information provided to test whether there is a significant relationship between the price of a house and the number of rooms at the 5% level.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 17

A 98% confidence interval for the mean of a population is to be constructed and must be accurate to within 0.3 unit. A preliminary sample standard deviation is 1.7. The smallest sample size n that provides the desired accuracy is

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 18

Identify the most appropriate test to use for the following situation: A national computer retailer believes that the average sales are greater for salespersons with a college degree. A random sample of 14 salespersons with a degree had an average weekly sale of $3542 last year, while 17 salespersons without a college degree averaged $3301 in weekly sales. The standard deviations were $468 and $642 respectively. Is there evidence to support the retailer's belief?

a) One sample t test
b) Matched pairs
c) Two sample t test
d) Two sample p test

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS