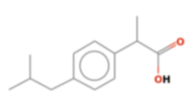
Exam II Solutions
Administered: Monday, October 12, 2020
22 points

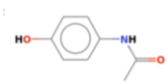
For each problem part: 0 points if not attempted or no work shown,
1 point for partial credit, if work is shown,
2 points for correct numerical value of solution

Problem 1. (16 points) Consider the following data for the enthalpy of fusion for two biochemicals.

Ibuprofen
$C_{13}H_{18}O_2$



Acetaminophen
$C_8H_9NO_2$



taken from the NIST Chemistry Webbook, http://webbook.nist.gov/chemistry/.

Enthalpy of fusion of ibuprofen

$\Delta_{fus}H$ (kJ/mol)	Temperature (K)	Method	Reference
39.5	350.4	DSC	Cilurzo, Alberti, et al., 2010
27.94	347.6	DSC	Hong, Hua, et al., 2010
26.6	346.4	DSC	Wassvik, Holmén, et al., 2006
26.65	348.	N/A	Gracin and Rasmuson, 2002
25.7	350.9	DSC	Li, Zell, et al., 1999

Enthalpy of fusion of acetaminophen

$\Delta_{fus}H$ (kJ/mol)	Temperature (K)	Method	Reference
27.6	443.2	DSC	Mota, Carneiro, et al., 2009
27.0	440.3	DSC	Vecchio and Tomassetti, 2009
26.49	441.9	AC,DSC	Xu, Sun, et al., 2006
26.2	443.	DSC	Romero, Bustamante, et al., 2004
26.02	441.2	N/A	Manzo and Ahumada, 1990

Perform the following tasks.
(a) Determine the sample mean of the enthalpy of fusion of ibuprofen.
(b) Determine the sample mean of the enthalpy of fusion of acetaminophen.
(c) Determine the sample variance of the enthalpy of fusion of ibuprofen.
(d) Determine the sample variance of the enthalpy of fusion of acetaminophen.
(e) Identify the appropriate distribution to describe the difference of means in this case?
(f) Determine the lower limit of a 95% confidence interval on the difference of means of the enthalpy of fusion.
(g) Determine the upper limit of a 95% confidence interval on the difference of means of the enthalpy of fusion.
(h) Explain your findings in language a non-statistician can understand.

Solution:

(a) Determine the sample mean of the enthalpy of fusion of ibuprofen.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{5} x_i = 29.278 \ \frac{kJ}{mol}$$

Based on this data set sample mean of the enthalpy of fusion of ibuprofen is 29.3 kJ/mol.

(b) Determine the sample mean of the enthalpy of fusion of acetaminophen.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{5} x_i = 26.662 \ \frac{kJ}{mol}$$

Based on this data set sample mean of the enthalpy of fusion of acetaminophen is 26.7 kJ/mol.

(c) Determine the sample variance of the enthalpy of fusion of ibuprofen.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = 33.28992 \left(\frac{kJ}{mol}\right)^2$$

Based on this data set the sample variance of the enthalpy of fusion of ibuprofen is 33.3 $(kJ/mol)^2$.

(d) Determine the sample variance of the enthalpy of fusion of acetaminophen.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = 0.41232 \left(\frac{kJ}{mol}\right)^2$$

Based on this data set the sample variance of the enthalpy of fusion of acetaminophen is 0.41 $(kJ/mol)^2$.

(e) Identify the appropriate distribution to describe the mean of the enthalpy of fusion in this case.

In this case we do not know the true population variance so the appropriate distribution of the difference of sample means is the **t** distribution.

(f) Determine the lower limit of a 98% confidence interval on the difference of means of the enthalpy of fusion.
(g) Determine the upper limit of a 98% confidence interval on the difference of means of the enthalpy of fusion.

$$C.I. = 1 - 2\alpha = 0.95$$

$$\alpha = \frac{1 - C.I.}{2} = \frac{1 - 0.95}{2} = 0.025$$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left[\left(\frac{s_1^2}{n_1}\right)^2 \Big/ (n_1 - 1)\right] + \left[\left(\frac{s_2^2}{n_2}\right)^2 \Big/ (n_2 - 1)\right]} \quad \text{if } \sigma_1 \neq \sigma_2$$

$$v = 4.08 \approx 4$$

The limits on the t-distribution are given by

```
>> t = icdf('t',0.025,4)
```

```
t = -2.7764
```

and for the upper limit

```
>> t = icdf('t',0.975,4)

t =  2.7764
```

We next insert all of these numbers into the equation for the confidence interval on the difference of means.

$$P\left[\left(\overline{X}_1-\overline{X}_2\right)+t_{\alpha}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\left(\mu_1-\mu_2\right)<\left(\overline{X}_1-\overline{X}_2\right)+t_{1-\alpha}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right]=1-2\alpha$$

$$P[-4.588<(\mu_1-\mu_2)<\ 9.828]=0.95$$

(h) Translate your conclusions from (d) and (e) into a sentence that a non-statistician can understand.

We are 95% confidence that the difference between the enthalpy of fusion of ibuprofen and acetaminophen lies within the range from -4.588 to 9.828 kJ/mol. While the sample mean of the enthalpy of fusion is larger for ibuprofen than acetaminophen, this confidence interval includes both positive and negative numbers, indicating a lack of confidence at this level that the enthalpy of fusion of ibuprofen is actually greater than that of acetaminophen.

Problem 2. (6 points)
We attend a (socially distanced and outdoor) concert featuring an accomplished guitarist, who plays a 12 string guitar. In the hands of this guitarist, who practices several hours every day, the lifetime of an individual guitar string follows the normal distribution with a mean of 3.7 months and a standard deviation of 0.5 months.
(a) What is the probability that a string lasts at least 4.0 months?
(b) What is the probability that this 12-string guitar has zero strings break (all 12 strings working) in 4.0 months?
(c) What is the probability that this 12-string guitar has more than 6 strings break (less than 6 strings working) in 4.0 months?

Solution:

(a) What is the probability that a string lasts at least 4.0 months?

$$P(x \geq 4.0) = 1 - P(x \leq 4.0)$$

We use the cdf function in MATLAB:

```
>> p = 1 - cdf('normal',4.0,3.7,0.5)

p =   0.274253117750074
```

There is a 27.4% chance that a string lasts at least 4.0 months.

(b) What is the probability that this 12-string guitar has zero strings break (all strings working) in 4.0 months?

This problem requires the binomial distribution. If we define the random variable x as the number of working strings, then x=12, n=12 and p, the probability of a single string still functioning after 4.0 months, defined in part a. We use the pdf function in Matlab because the binomial distribution is discrete and a discrete pdf gives the probability that x = a given value.

$$P(x = 12) = \text{pdf}('\text{binomial}', \text{x} = 12, \text{n} = 12, \text{p} = 0.274253117750074)$$

```
>> p = pdf('binomial',12,12,0.274253117750074)

p =     1.810584844491894e-07
```

There is almost no chance that none of the strings have broken after 4.0 months.

Alternatively, if we define the random variable x as the number of broken strings, then x=0, n=12 and p, the probability of a single string still functioning after 4.0 months, defined as one minus the value from part a. We use the pdf function in Matlab because the binomial distribution is discrete and a discrete pdf gives the probability that x equals a given value.

$$P(x = 12) = \text{pdf}('\text{binomial}', \text{x} = 12, \text{n} = 12, \text{p} = 1 - 0.274253117750074)$$

```
>> p = pdf('binomial',0,12,1-0.274253117750074)

p =     1.810584844491894e-07
```

This is the same result as above. The definition of success (working string or broken string) is arbitrary, but the probability must correspond to the definition of success.

(c) What is the probability that this 12-string guitar has more than 6 strings break (less than 6 strings working) in 4.0 months?

This problem requires the binomial distribution with x<10, n=12 and p, the probability of a single string still functioning after 4.0 months, defined in part a. The cdf function calculates the probability that a random variable is less than or equal to a number, so we need to convert the less than expression to one containing less than or equal to.

$$P(x < 6) = P(x \leq 5)$$

$$P(x \leq 5) = \text{cdf}('\text{binomial}', \text{x} = 5, \text{n} = 12, \text{p} = 0.274253117750074)$$

```
>> p = cdf('binomial',5,12,0.274253117750074)

p =   0.918718390172291
```

There is a 91.9% chance that the guitar has less than six strings working after 4.0 months.