Exam II Solutions
Administered: Monday, October 11, 2021
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 temperature of fusion for two alcohols.

1-octanol
$C_8H_{18}O$

1-decanol
$C_{10}H_{22}O$

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

Temperature of Fusion of 1-octanol

T_{fus} (K)	Reference	Comment
258.35	Liu, Pusicha, et al., 1991	Uncertainty assigned by TRC = 0.05 K; TRC
258.03	Plesnar, Gierycz, et al., 1988	Uncertainty assigned by TRC = 0.08 K; TRC
256.2	Davies and Kybett, 1965	Uncertainty assigned by TRC = 0.5 K; TRC
256.25	Costello and Bowden, 1958	Uncertainty assigned by TRC = 0.4 K; TRC
256.4	Tschamler, Richter, et al., 1949	Uncertainty assigned by TRC = 0.5 K; TRC
256.9	Timmermans, 1922	Uncertainty assigned by TRC = 0.4 K; TRC

Temperature of fusion of 1-decanol

T_{fus} (K)	Reference	Comment
279.6	Davies and Kybett, 1965	Uncertainty assigned by TRC = 0.5 K; *TRC*
280.05	Costello and Bowden, 1958	Uncertainty assigned by TRC = 0.4 K; *TRC*
277.	Badin, 1943	Uncertainty assigned by TRC = 3. K; *TRC*

Perform the following tasks.
(a) Determine the sample mean of the temperature of fusion of 1-octanol.
(b) Determine the sample mean of the temperature of fusion of 1-decanol.
(c) Determine the sample variance of the temperature of fusion of 1-octanol.
(d) Determine the sample variance of the temperature of fusion of 1-decanol.
(e) Identify the appropriate distribution to describe the difference of means in this case?
(f) Determine the lower limit of a 98% confidence interval on the difference of means of the temperature of fusion.
(g) Determine the upper limit of a 98% confidence interval on the difference of means of the temperature of fusion.
(h) Explain your findings in language a non-statistician can understand.

Solution:

(a) Determine the sample mean of the temperature of fusion of 1-octanol.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{6} x_i = 257.02\ K$$

Based on this data set sample mean of the temperature of fusion of 1-octanol is 257.0 K.

(b) Determine the sample mean of the temperature of fusion of 1-decanol.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{3} x_i = 278.88\ K$$

Based on this data set sample mean of the temperature of fusion of 1-decanol is 278.9 K.

(c) Determine the sample variance of the temperature of fusion of 1-octanol.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = 0.8906(K)^2$$

Based on this data set the sample variance of the temperature of fusion of 1-octanol is 0.891 K^2.

(d) Determine the sample variance of the temperature of fusion of 1-decanol.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = 2.7108(K)^2$$

Based on this data set the sample variance of the temperature of fusion of 1-decanol is 2.71 K^2.

(e) Identify the appropriate distribution to describe the mean of the temperature 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 temperature of fusion.
(g) Determine the upper limit of a 98% confidence interval on the difference of means of the temperature of fusion.

$$C.I. = 1 - 2\alpha = 0.98$$

$$\alpha = \frac{1 - C.I.}{2} = \frac{1 - 0.98}{2} = 0.01$$

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

$$v = 2.68$$

The limits on the t-distribution are given by

```
>> t = icdf('t',0.01,2.68)
```

```
t = -5.0040
```

and for the upper limit

```
>> t = icdf('t',0.99,2.68)

t =  5.0040
```

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

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

$$P[-26.9942 < (\mu_1 - \mu_2) < -16.7291] = 0.98$$

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

We are 98% confidence that the difference between the temperature of fusion of 1-octanol and 1-decanol lies within the range from -26.99 K to -16.73 K. In other words, we are 98% confident that the temperature of fusion of 1-decanol is between 26.99 and 16.73 K higher than the temperature of fusion of 1-octanol.

Problem 2. (6 points)
An apartment complex with 30 units has a historical database that describes the length of rental of its residents. The length of rental is described by the gamma distribution with parameters $\alpha = 2$ and $\beta = 3$ years.
(a) What is the probability that a resident stays in the apartment complex for 5 years or longer?
(b) Assuming full occupancy, what is the probability that exactly half of the units have the same occupants after 5 years?
(c) Assuming full occupancy, what is the probability that at least half of the units have the same occupants after 5 years?

Solution:

(a) What is the probability that a resident stays in the apartment complex for 5 years or longer?

$$P(x \geq 5.0) = 1 - P(x \leq 5.0)$$

We use the cdf function in MATLAB:

```
>> p = 1 - cdf('gamma',5,2,3)

p =   0.503668274233498
```

There is a 50.4% chance that a resident stays in the apartment complex for 5 years or longer.

(b) Assuming full occupancy, what is the probability that exactly half of the units have the same occupants after 5 years?

This problem requires the binomial distribution. If we define the random variable x as the number of units with the same occupant for the past five years, then x=15, n=30 and p is the probability of having the same resident for five years or longer, 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.

$$P(x = 12) = \mathrm{pdf}('\mathrm{binomial}', \mathrm{x} = 15, \mathrm{n} = 30, \mathrm{p} = 0.503668274233498)$$

```
>> p = pdf('binomial',15,30,0.503668274233498)

p =     0.144347855168778
```

There is a 14.4% probability that half the units have the same residents after five years.

(c) Assuming full occupancy, what is the probability that at least half of the units have the same occupants after 5 years?

This problem requires the binomial distribution with $x \geq 15$, n=30 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.

$$P(x \geq 15) = 1 - P(x \leq 14)$$

$$P(x \leq 5) = \mathrm{cdf}('\mathrm{binomial}', \mathrm{x} = 14, \mathrm{n} = 30, \mathrm{p} = 0.503668274233498)$$

```
>> p = 1 - cdf('binomial',14,30,0.503668274233498)

p =   0.588067991477964
```

There is a 58.8% chance that at least half of the units have the same occupants after 5 years.