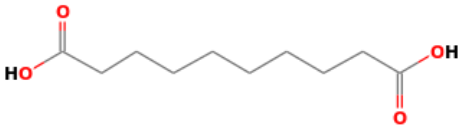
Exam II Solutions
Administered: Monday, October 9, 2026
24 points

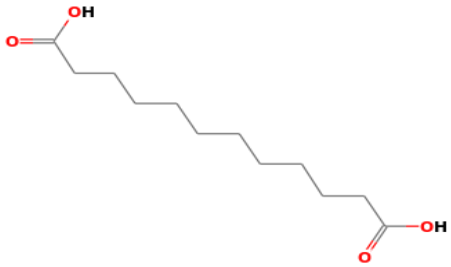
For each problem part: 0 points if not attempted or no work shown,
1 point for partial credit, if work is shown,
2 points for correct numerical value of solution

Problem 1. (16 points)
Consider the following data for the enthalpy of fusion for two alkanes of different lengths with a carboxylic acid group on each end.

decanedioic acid
$C_{10}H_{18}O_4$



dodecanedioic acid
$C_{12}H_{22}O_4$



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

Enthalpy of fusion of decanedioic acid

$\Delta_{fus}H$ (kJ/mol)	Temperature (K)	Method	Reference
40.807	404.0	N/A	Cingolani and Berchiesi, 1974
46.9	405.7	N/A	Ventolà, Bayés, et al., 2008
45.3	405.6	DSC	Xia, Zhang, et al., 2008
46.6	403.9	N/A	Roux, Temprado, et al., 2005
40.8	404.	N/A	Domalski and Hearing, 1996

Enthalpy of fusion of dodecanedioic acid

$\Delta_{fus}H$ (kJ/mol)	Temperature (K)	Method	Reference
50.564	402.5	N/A	Cingolani and Berchiesi, 1974
52.5	401.6	N/A	Ventolà, Bayés, et al., 2008
49.8	400.3	DSC	Roux, Temprado, et al., 2005
50.57	402.5	N/A	Domalski and Hearing, 1996

Perform the following tasks.
(a) Determine the sample mean of the enthalpy of fusion of decanedioic acid.
(b) Determine the sample mean of the enthalpy of fusion of dodecanedioic acid.
(c) Determine the sample variance of the enthalpy of fusion of decanedioic acid.
(d) Determine the sample variance of the enthalpy of fusion of dodecanedioic acid.
(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 decanedioic acid.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 44.08\ \frac{kJ}{mol}$$

Based on this data set sample mean of the enthalpy of fusion of decanedioic acid is 44.1 kJ/mol.

(b) Determine the sample mean of the enthalpy of fusion of dodecanedioic acid.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 50.86\ \frac{kJ}{mol}$$

Based on this data set sample mean of the enthalpy of fusion of dodecanedioic acid is 50.9 kJ/mol.

(c) Determine the sample variance of the enthalpy of fusion of decanedioic acid.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = 9.32\left(\frac{kJ}{mol}\right)^2$$

Based on this data set the sample variance of the enthalpy of fusion of decanedioic acid is 9.32 $(kJ/mol)^2$.

(d) Determine the sample variance of the enthalpy of fusion of dodecanedioic acid.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} = 1.33\left(\frac{kJ}{mol}\right)^2$$

Based on this data set the sample variance of the enthalpy of fusion of dodecanedioic acid is 1.33 $(kJ/mol)^2$.

(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 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.

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

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

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

$$v = 5.327$$

The limits on the t-distribution are determined using the Python code given below.

`tlo = -2.5237953907390587` and `thi = 2.5237953907390587`

note: If you use Excel to calculate the t limits, it rounds the degrees of freedom to the nearest integer, in this case 5, which changes the limits on the t distribution to

`tlo = -2.570581836` and `thi = 2.570581836`

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

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

$$P[-10.52 < (\mu_1 - \mu_2) < \text{-}3.03] = 0.95$$

If you rounded the degrees of freedom to an integer, your corresponding confidence interval is

$$P[-10.59 < (\mu_1 - \mu_2) < \text{-}2.97] = 0.95$$

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

We are 95% confidence that the difference between the enthalpy of fusion of decanedioic acid and dodecanedioic acid lies within the range from -10.52 to -3.03 kJ/mol. The sample mean of the enthalpy of fusion is larger for dodecanedioic acid than decanedioic acid, so these negative limits make sense.

*** xm02_p01.py code start ***

```
import numpy as np
import statistics
from scipy.stats import t

# Step 1.  Input Data
print("Step 1.  Input Data")
datamat1 = np.array([40.807, 46.9, 45.3, 46.6, 40.8])
datamat2 = np.array([50.564, 52.5, 49.8, 50.57])
ndata1 = datamat1.shape[0]
ndata2 = datamat2.shape[0]
print(f"The first data set has a total of {ndata1} data points.")
print(f"The second data set has a total of {ndata2} data points.")

# Step 2.  determine statistics
print("Step 2.  determine statistics")
sample_mean1 = statistics.mean(datamat1)
sample_mean2 = statistics.mean(datamat2)
print(f"(a) The first data set has a mean of {sample_mean1}.")
print(f"(b) The second data set has a mean of {sample_mean2}.")
sample_var1 = statistics.variance(datamat1)
sample_var2 = statistics.variance(datamat2)
print(f"(c) The first data set has a sample variance of {sample_var1}.")
print(f"(c) The second data set has a sample variance of {sample_var2}.")
```

```
# Step 3.  create confidence intervals
print("Step 4. create confidence intervals")
# confidence interval
confint = 0.95
diffmeans = sample_mean1 - sample_mean2
fac = np.sqrt( sample_var1/ndata1 + sample_var2/ndata2 )
confint = 0.95
alpha = 0.5*(1.0 - confint)
alpha = 0.5*(1.0 - confint)
plo = alpha
phi = 1.0 - alpha
dof_numer = (sample_var1/ndata1 + sample_var2/ndata2)**2
dof_denom1 = (sample_var1/ndata1)**2 / (ndata1 - 1)
dof_denom2 = (sample_var2/ndata2)**2 / (ndata2 - 1)
dof_denom = dof_denom1 + dof_denom2
dof = dof_numer/dof_denom
print(f"dof is {dof}.")
tlo = t.ppf(plo, df=dof)
print(f"tlo is {tlo}.")
thi = t.ppf(phi, df=dof)
print(f"thi is {thi}.")
dmulo = diffmeans + tlo*fac
dmuhi = diffmeans + thi*fac
print(f"(f) & (g) P( {dmulo} < dmu < {dmuhi} ) is {confint}")
```

*** xm02_p01.py code end ***

This code is executed by typing:

```
run xm02_p01.py
```

which yields the following output

```
Step 1.  Input Data
The first data set has a total of 5 data points.
The second data set has a total of 4 data points.
Step 2.  determine statistics
(a) The first data set has a mean of 44.0814.
(b) The second data set has a mean of 50.8585.
(c) The first data set has a sample variance of 9.315529799999998.
(c) The second data set has a sample variance of 1.3283023333333353.
Step 4. create confidence intervals
dof is 5.327319921340232.
tlo is -2.5237953907390587.
thi is 2.5237953907390587.
(f) & (g) P( -10.516391860195277 < dmu < -3.037808139804717 ) is 0.95
```

Problem 2. (8 points)

Consider a battery with a lifetime described by the normal distribution with a population mean of 5.0 years and a population standard deviation of 0.5 years.

(a) What is the probability that a battery lasts at least 6.0 years?

(b) What PDF would describe the probability that x batteries haven't failed before 6.0 years?

(c) What is the probability that a device with a redundant power system running on three batteries has no batteries working in 6.0 years?

(d) What is the probability that a device with a redundant power system running on three batteries has at least one battery working in 6.0 years?

Solution:

(a) What is the probability that a battery lasts at least 6.0 years?

We are being asked to determine the probability that the life time of a single battery is greater than 6.0 years minutes. This can be related to the CDF as

$$P(t > 6) = 1 - P(t \leq 6)$$

The following python code can evaluate this result

*** xm02_p02 code start ***

```
from scipy.stats import norm

# part (a)

mean = 5.0
stdev = 0.5
x = 6.0
pcdf = norm.cdf(x, loc=mean, scale=stdev)
parta = 1.0 - pcdf
print(f"(a) prob = {parta}")
```

*** hw01_p02 code end ***

Running this script yields the following result,

```
run xm02_p02.py
(a) prob = 0.02275013194817921
```

The probability that a battery has a lifetime greater than 6.0 years is 0.0228.

(b) What PDF would describe the probability that x batteries haven't failed before 6.0 years?

This calls for the binomial probability because the batteries are independent and all have the same probability of not failing before 6.0 years.

x = random variable = number of batteries functioning after 6.0 years
n = total number of samples = 3
p = probability that an individual battery hasn't failed before 6.0 years = answer to part (a)

(c) What is the probability that a device with a redundant power system running on three batteries has no batteries working in 6.0 years?

We are being asked for the probability that zero batteries are functioning after 6.0 years. This can be related to the PDF as

$$P(x = 0) = b(0; 3, 0.0228)$$

We added a bit more code to python script used to solve part (a) can evaluate this result

*** hw01_p02 code start ***

```
from scipy.stats import norm
from scipy.stats import binom

# part (a)
mean = 5.0
stdev = 0.5
x = 6.0
pcdf = norm.cdf(x, loc=mean, scale=stdev)
parta = 1.0 - pcdf
print(f"(a) prob = {parta}")

# part (c)
n = 3
x = 0
partc = binom.pmf(x, n, parta)
print(f"(c) prob = {partc}")
```

*** hw01_p02 code end ***

Running this script yields the following result,

```
run xm02_p02.py
(a) prob = 0.02275013194817921
(c) prob = 0.9332905349146907
```

$$P(x = 0) = b(0; 3, 0.0228) = 0.9333$$

The probability that zero batteries still function after 6.0 years is 0.9333.

(d) What is the probability that a device with a redundant power system running on three batteries has at least one battery working in 6.0 years?

We are being asked for the probability that zero batteries are functioning after 6.0 years. This can be related to the CDF as

$$\begin{aligned} P(x \geq 1) &= 1 - P(x \leq 0) \\ &= 1 - B(0; 3, 0.0228) \end{aligned}$$

We added a bit more code to python script used to solve part (a) can evaluate this result

*** hw01_p02 code start ***

```
from scipy.stats import norm
from scipy.stats import binom

# part (a)
mean = 5.0
stdev = 0.5
x = 6.0
pcdf = norm.cdf(x, loc=mean, scale=stdev)
parta = 1.0 - pcdf
```

```
print(f"(a) prob = {parta}")

# part (c)
n = 3
x = 0
partc = binom.pmf(x, n, parta)
print(f"(c) prob = {partc}")

# part (d)
n = 3
x = 0
partd = 1.0 - binom.cdf(x, n, parta)
print(f"(d) prob = {partd}")
```

*** hw01_p02 code end ***

Running this script yields the following result,

```
run xm02_p02.py
(a) prob = 0.02275013194817921
(c) prob = 0.9332905349146907
(d) prob = 0.0667094650853094
```

$$P(x \geq 1) = 1 - P(x \leq 0) = 0.0667$$

The probability that at least one battery still functions after 6.0 years is 0.0667.