Exam II Solutions
Administered: Monday, October 10, 2025
26 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 with work

Problem 1. (20 points)
Consider the following data is provided for the enthalpy of fusion of octane, taken from the NIST Chemistry Webbook, http://webbook.nist.gov/chemistry/.

Enthalpy of fusion of Octane (C_8H_{18})

$\Delta_{fus}H$ (kJ/mol)	Temperature (K)	Reference
20.740	216.38	Finke, Gross, et al., 1954
21.8	216.6	Mondieig, Rajabalee, et al., 2004
20.74	216.4	Domalski and Hearing, 1996
20.652	215.8	Huffman, Parks, et al., 1931
20.092	215.6	Parks, Huffman, et al., 1930

Perform the following tasks.
(a) Determine the sample mean of the enthalpy of fusion of octane.
(b) Determine the sample variance of the enthalpy of fusion of octane.
(c) Determine the sample standard deviation of the enthalpy of fusion of octane.
(d) Identify the appropriate distribution to describe the mean of the enthalpy of fusion of octane in this case.
(e) Determine the lower limit of a 98% confidence interval on the mean of the enthalpy of fusion of octane.
(f) Determine the upper limit of a 98% confidence interval on the mean of the enthalpy of fusion of octane.
(g) Identify the appropriate distribution to describe the variance of the enthalpy of fusion of octane in this case.
(h) Determine the lower limit of a 98% confidence interval on the variance of the enthalpy of fusion of octane.
(i) Determine the upper limit of a 98% confidence interval on the variance of the enthalpy of fusion of octane.
(j) Explain your findings in parts (a) through (i) in language a non-statistician can understand.

Solution:

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

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{5} x_i = \frac{20.740 + 21.8 + 20.74 + 20.652 + 20.092}{5} = 20.8048\ \frac{kJ}{mol}$$

Based on this data set sample mean of the enthalpy of fusion of octane is 20.80 kJ/mol.

(b) Determine the sample variance of the enthalpy of fusion of octane.

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} = \frac{5(2165.729) - (104.024)^2}{5(5-1)} = 0.382563\left(\frac{kJ}{mol}\right)^2$$

Based on this data set the sample variance of the enthalpy of fusion of octane is 0.383 $(kJ/mol)^2$.

(c) Determine the sample standard deviation of the enthalpy of fusion of octane.

$$s = \sqrt{s^2} = \sqrt{0.382563} = 0.618517 \frac{kJ}{mol}$$

Based on this data set the sample standard deviation of the enthalpy of fusion of octane is 0.619 kJ/mol.

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

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

(e) Determine the lower limit of a 98% confidence interval on the mean of the enthalpy of fusion of octane.
(f) Determine the upper limit of a 98% confidence interval on the mean of the enthalpy of fusion of octane.

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

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

$$v = n - 1 = 4$$

The limits on the t-distribution are given by

$$t_{0.01} = -3.7469$$

and for the upper limit

$$t_{0.99} = 3.7469$$

We next insert all of these numbers into the equation for the confidence interval on the sample mean.

$$P\left[\bar{X} + t_\alpha \frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{1-\alpha} \frac{s}{\sqrt{n}}\right] = 1 - 2\alpha$$

$$P\left[20.80\text{-}3.7469 \frac{0.619}{\sqrt{5}} < \mu < 20.80 + 3.7469 \frac{0.619}{\sqrt{5}}\right] = 0.98$$

$$P[19.768 < \mu < 21.841] = 0.98$$

Based on this data set we are 98% confident that the true population mean of the enthalpy of fusion of octane lies between 19.77 and 21.84 kJ/mol.

(g) Identify the appropriate distribution to describe the variance of the enthalpy of fusion of octane in this case.

The appropriate distribution of the sample variance is the $\boldsymbol{\chi^2}$ (chi-squared) distribution.

(h) Determine the lower limit of a 98% confidence interval on the variance of the enthalpy of fusion of octane.
(i) Determine the upper limit of a 98% confidence interval on the variance of the enthalpy of fusion of octane.

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

$$\alpha = \frac{1 - C.I.}{2} = \frac{1 - 0.98}{2} = 0.001$$

$$v = n - 1 = 4$$

The limits on the χ^2-distribution are given by

$$\chi^2_{0.01} = 0.2971$$

and for the upper limit

$$\chi^2_{0.99} = 13.2767$$

We next insert all of these numbers into the equation for the confidence interval on the sample variance.

$$P\left[\frac{(n-1)s^2}{\chi^2_{1-\alpha}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha}}\right] = 1 - 2\alpha$$

$$P\left[\frac{(5-1)0.3826}{13.2767} < \sigma^2 < \frac{(5-1)0.3826}{0.2971}\right] = 0.98$$

$$P[0.115 < \sigma^2 < 5.150] = 0.98$$

Based on this data set we are 98% confident that the true population variances lies between 0.115 and 5.150 $(\text{kJ/mol})^2$.

(j) Explain your findings in parts (a) through (i) in language a non-statistician can understand.

Based on all 5 data points in the literature, we can make the following observations. The mean of the enthalpy of fusion of octane is 20.80 kJ/mol with a 98% confident interval in the range from 19.77 and 21.84 kJ/mol. The variance of the enthalpy of fusion of octane is 0.3826 $(\text{kJ/mol})^2$ with a 98% confident interval in the range from 0.115 and 5.150 $(\text{kJ/mol})^2$.

I modified the following python script to solve Problem 1. The script has the ability to evaluate outlier, but I turned that function off to use all the data. The script calls pre-existing routines as discussed in class for evaluating the sample statistics and the confidence intervals.

*** hw02_p01.py code start ***

```
# Exam 2 Problem 1
# Author: David Keffer dkeffer@utk.edu
# Date begun: October 3, 2025
# Last updated: October 3, 2025

import numpy as np
from find_outliers_funk import find_outliers_funk
from get_sample_stats_wo_outliers import get_sample_stats_wo_outliers
from confint_meanvarunknown_funk import confint_meanvarunknown_funk
from confint_var_funk import confint_var_funk

# Step 1.  Input Data
print("Step 1.  Input Data")

datamat = np.array([
[20.740],
[21.8  ],
[20.74 ],
[20.652],
[20.092],
])

ndata, nprop = datamat.shape
print(f"The data set has a total of {ndata} data points.")

# Step 2. identify outliers
# if outlier_check is 1, check for outliers
# if outlier_check is 0, use all data
outlier_check = 0;
if outlier_check == 1:
    print("Step 2. identify outliers")
    ngood, good = find_outliers_funk(datamat)
    print(f"The first data set has {ngood} data points.")
    for i in range(ndata):
        if good[i] == 0:
            print(f"data {i} ({datamat[i]}) is an outlier")
else:
    print("Using all data for statistics.")
    good = np.ones(ndata)
    ngood = ndata

# Step 3.  determine statistics
print("Step 3.  determine statistics")
mean_samp, var_samp, sd_samp = get_sample_stats_wo_outliers(datamat, good, ngood)

# Step 4.  create confidence intervals
print("Step 4. create confidence intervals")

# confidence interval
confint = 0.98

# Step 4.1. create confidence interval on mean
print("   Step 4.1. create confidence interval on mean")
ci_mulo, ci_muhi = confint_meanvarunknown_funk(mean_samp, sd_samp, ngood, confint)

# Step 4.2. create confidence interval on variance
print("   Step 4.3. create confidence interval on variance")
ci_siglo, ci_sighi = confint_var_funk(var_samp, ngood, confint)
```

*** hw02_p01.py code end ***

Problem 2. (6 points)

The mechanical properties of four dogbone samples are characterized in four independent devices. A constant tensile stress is applied over an extended period of time. At a given temperature, the time to failure of an individual dogbone is known to follow the normal distribution with a mean lifetime of 30 minutes and a standard deviation of 3 minutes.

(a) What is the probability that an individual dogbone sample hasn't failed before 25 minutes?
(b) What PDF would describe the probability that x of the 4 dogbones haven't failed before 25 minutes?
(c) What is the probability that at least two dogbones haven't yielded before 25 minutes?

Solution:

(a) What is the probability that an individual dogbone sample hasn't failed before 25 minutes?

We are being asked to determine the probability that the failure time of a dogbone is greater than 25 minutes. This can be related to the CDF as

$$P(t > 25) = 1 - P(t \leq 25)$$

The following python code can evaluate this result

*** hw01_p02 code start ***

```
from scipy.stats import norm

# part (a)

mean = 30.0
stdev = 3.0
x = 25.0
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
(a) prob = 0.9522096477271853
```

The probability that a dogbone hasn't failed in 25 minutes is 0.952.

(b) What PDF would describe the probability that x dogbones haven't failed before 25 minutes?

This calls for the binomial probability because the samples are independent and all have the same probability of not failing before 25 minutes.

x = random variable = number of dogbones functioning after 25 minutes
n = total number of samples= 4
p = probability that an individual dogbone sample hasn't yielded before 25 minutes = answer to part (a)

(c) What is the probability that at least two dogbones haven't yielded before 25 minutes?

We are being asked for the probability that the number of dogbones that haven't failed is greater than or equal to two. This can be related to the CDF as

$$P(x \geq 2) = 1 - P(x \leq 1)$$

$$P(x \geq 2) = 1 - B(1; 4, 0.9522)$$

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

*** hw01_p02 code start ***

```
import numpy as np
from scipy.stats import norm
from scipy.stats import binom

# part (a)

mean = 30.0
stdev = 3.0
x = 25.0
pcdf = norm.cdf(x, loc=mean, scale=stdev)
parta = 1.0 - pcdf
print(f"(a) prob = {parta}")

# part (c)
n = 4
x = 1
pcdf = binom.cdf(x, n, parta)
partc = 1.0 - pcdf
print(f"(c) prob = {partc}")
```

*** hw01_p02 code end ***

Running this script yields the following result,

```
run xm02_p02
(a) prob = 0.9522096477271853
(c) prob = 0.9995790519019148
```

$$P(x \geq 2) = 1 - B(1; 4, 0.9522) = 0.9996$$

The probability that at least two dogbones haven't failed in 25 minutes is 0.9996.