**Q1: Given that a student has an engineering background, what is the probability they score 70 or less in Component 1?**

Assume that the scores of students with an engineering background in Component 1 are normally distributed with:

- Mean ($\mu$) = 75
- Standard Deviation ($\sigma$) = 3

We want to find the probability of scoring 70 or less.
Solution:

1. Calculate the z-score: $z=(70-75)/3=-1.67$
2. Find the probability:
   - Use a standard normal distribution table or a calculator: $P(Z \leq -1.67) \approx 0.0475$
   - So, the probability is approximately 4.75%.

**Q2: In the given data set, what fraction of students with an engineering background have scored 70 or less in component 1?**

We can use the normal distribution to calculate the probability of a student scoring 70 or less in component 1, given that they have an engineering background. Let X be the random variable representing the component 1 score of a student with an engineering background. We know that X follows a normal distribution with mean 75and standard deviation 3. We can standardize X as follows:

$Z = (X - 75) / 3$
The probability that a student scores 70 or less is equivalent to the probability that Z is less than or equal to $(70-75)/3 = -5/3$. We can find this probability using a standard normal distribution table or a calculator:

$P(Z =-5/3) \approx 0.0918$
Therefore, about 9.18% of students with an engineering background scored 70 or less in component.

**Q3. Given the distributions, what is the expected value of the class score in Component 1? (Do not use the data set for this question)**

Explanation:

- Let $P(E)=0.6$, $P(C)=0.3$, $P(O)=0.1$
- Given:
  - $E[X|E] =75$
  - $E[X|C] =72$
  - $E[X|O] =78$

Solution:

$E[X]=(75\times0.6)+(72\times0.3)+(78\times0.1)=74.$

**Q4. In the given data set, what is the average score by students in Component 1?**

Solution (Using the Dataset):

1. Calculate the average score:
   - Use Excel formula:
     =AVERAGE(D:D)

Answer: 75.47.

**Q5. Given that a student scored 80 or more in Component 1, what is the probability that this student is neither from an engineering background nor a commerce background?** (Do not use the dataset for this question)

This is the case of conditional probability

P(O/S >= 80),finding probability of student being from other background given than student scored 80 or more in component 1.

P(O/S>= 80 ) = P(O ∩ S >= 80) / P(S>= 80)
P(O ∩ S >= 80) can be rewritten as P(O) * P(S>=30/O)
So finally ; P(O/S>= 80) = P(O) * P(S>=30/O)/ P(S>= 80) –- (a)

Let P(S>= 80) = P(B) :- P(B) = P(B/E)*P(E) + P(B/C)*P(C) + P(B/O)*P(O) –(b)
Solve : -
P(B/E) = 80-75/3 = 5/3 = 1.67
P(s>=80/E) = 1- P(s <= 80/E) = 1- 0.9525 = 0.0475
P(B/E)*P(E) = 0.0475*0.6 = 0.0285 –(1)

P(B/C) = 80-76/5 = ⅘ = 0.8
P(s>=80/C) = 1- P(s <= 80/C) = 1-0.7881 = 0.2119 P(B/C)*P(C) = 0.2119 * 0.3 = 0.06357 – (2)

P(B/O) = 80-85/4 = -5/4 = -1.25 P(s>=80/O) = 1- P(s <= 80/O) = 0.89435
P(B/O)*P(O) = 0.89435 *0.1 = 0.089435– (3)

Substituting values 1,2,3 in equation b
P(S>= 80) = 0.0285 + 0.06357 + 0.089435 =0.181505 –(c)

Substituting value in (a)
P(O/S>= 80) = (3)/(c) = 0.089435/0.181505= 0.4927 or 49.27%

So the probability student being from other background given than student scored 80 or more in component 1 is 49.27%

**Q6. What percentage of the students who have scored over 80 in component 1 are neither from an engineering background nor a commerce background?**

Total students who scored over 80 in component 1 = 29 – (1)
Students who scored over 80 in component 1 and are from other background = 11 – (2)
Percentage = (2)/(1) = 11/29 = 38%

**Q7. The final score obtained by a student is the average of the scores in the three components. Draw a sample of the students by choosing students with serial numbers 1, 11, 21, ... 291. Assume this to be a random sample.**
**a) Based on this sample, what is a point estimate of the mean score of students taking this course?**
**b) Based on this sample, what is a 95% confidence interval of the mean scores of students taking this course?**

a) Given final score is average of scores in three components Point estimate is the average of sample values, which in this case came out to 74.4889
b) Sample size = 30 (1,11,21,31,41,51….291)
   Sample standard deviation = 5.7062
   95% confidence interval means z value = 1.96 (0.05 is leftover probability that is outside the interval)
   Confidence interval = x ± (Z×SEM)
   = 74.4889 ± (1.96 x (5.7062/√30)) = 74.4889 ± 2.04
   Interval = [72.4489,76.5289]

**Q8. Suppose that the data set here represents a sample from a population of similar students. Based on the data set, would you conclude that students with engineering backgrounds have an average score of 75 in component 1? (Use α=0.05.)**

From the dataset we get the following details:-
1) Total engineering student = 184
2) Sample average = 74.8
3) Sample standard deviation = 2.8

Hypothesis: - Null hypothesis = μE = 75 Alternate hypothesis = μE <> 75

For 1 sample test , t test :-
t= sample mean - hypothesized mean / (s/√n) t= 74.8 - 75 / (2.8/√184)
t= -0.2/0.2064
t stats= -0.969
t critical for two tail test with df = 183 and α=0.05 is +/- 1.973
We know that ; t stats > t critical = reject the null hypothesis t stats <= t critical then we fail to reject the null hypothesis In our case, t stats lies between [-1.973 , + 1.973] so we fail to reject the null hypothesis.

So we conclude that there is sufficient evidence to conclude that the average score of engineering students is 75 at the 5% significance level. Or we conclude with 95% confidence that average score of engineering students is 75

**Q9. Suppose we choose two random samples of 30 students each from the class. The first sample contains students who have commerce backgrounds, and the second sample contains students who have engineering backgrounds. The average of the scores of the first sample is 75.8333, with a sample standard deviation of 5.7813. The average of the scores of the second sample is 74.7444, with a sample standard deviation of 3.4416. Based on these samples, would you conclude that students with commerce backgrounds scored better than students with engineering backgrounds? (Use α=0.05α=0.05.)**

Given that one sample is of commerce and second sample is of engineering
x1 = 75.8333, s1=5.7813 and x2 = 74.7444, s2 = 3.4416
Hypothesis :-
1) Null hypothesis => u1 = u2 (no difference in average score of commerce and engineering background)
2) Alternate hypothesis => u1 > u2

For checking whether two samples have equal variance: -

- Commerce Students:(s1)^2 = 33.424 – (1)
- Engineering Students: (s2)^2 = 11.845 –-(2)
- The variance ratio is: (1)/(2) = 33.424/11.845 = 2.82

The variance ratio is less than 4:1, so we can assume that they have equal variance

This is the case of independent two samples with equal variance.
Df = n1+n2-2 = 30+30-2 = 58 and α=0.05
The formula for independent two sample = t= x1-x2/(s1^2/n1 + s2^2/n2)^½
t= 75.8333-74.7444/(33.424/30+11.845/30)^½
t= 1.0889/1.228 = 0.886

P value (t stat = 0.886, df = 58, test = one tailed and α=0.05) = 0.18964
P value > 0.05 , we fail to reject the null hypothesis.
Conclusion: - With α=0.05 significance we conclude that the average scores are not significantly different of commerce and engineering background. In simpler terms, we can't say commerce students score better than engineering students.

**Q10. Based on the average of all three components, grades are awarded to students. The rulebook says that no more than 20% of the students can score the highest grade: A, and together, no more than 60% of the students can score the highest and second highest grade i.e., A and B. Students who score neither A nor B get a C grade. Now, the instructor wants to give as many A and B grades as the rule book permits. In this context, answer the following questions:**

**a) What is the average score obtained by students with A grades?**

**b) What is the average score obtained by students with B grades?**

**c) What fraction of students with an engineering background have scored A grades?**

**d) Among students who scored A grade, what fraction had engineering backgrounds?**

a) It is found that for falling in 20% (80th percentile) category student total score should be above = 79.267

Using percentile formula in the dataset,60 students scored above 79.267
Average score of 20% student found out to be = 82.41667

b) A grade +B grade = 60%(40th percentile)
B grade marks should be > 73.9 and < = 79.267 {40th percentile = 73.8 (score greater than this)}
So average score of B grade scoring student found out to be 76.21667

c) Total engineering students = 184
Total engineering students who scored A grade = 20
Fraction = 20/184 = 0.10870 i.e 10.87%

d) Total students who scored A grade = 60
Total engineering students who scored A grade = 20
Fraction = 20/60 = 0.33334 i.e 34%