1. Alice obtained a score of 26 on her ENGG 2780 exam. Her five friends' scores are 29, 27, 19, 24, and 17. Assume the scores are independent samples of a normal random variable.

   (a) Can Alice conclude that she did better than average with 75% confidence?

   Solution: We assume the data are Normal(μ, σ) samples for unknown μ. The null hypothesis is that $\mu = 26$ (Alice did average)[1] and the alternative hypothesis is $\mu \leq 26$. The test should accept when $T < t$ for a suitable threshold t, where $T = (\overline{X} - 26)/(S/\sqrt{5})$. Under the null hypothesis t is a $t(4)$ random variable so t should be set to satisfy $\mathrm{P}(t(4) < t) = 0.25$, which gives $t \approx -0.741$. Plugging in $\overline{X} = 23.2$ and $S \approx 5.12$ we obtain $T \approx -1.22$, so $T < t$ and Alice can conclude that she did better than average.

   (b) What is the p-value for the outcome in part (a)?

   Solution: The p-value is the smallest probability of a false positive assuming t is set just small enough so that the test accepts given the outcome. This is the probability that a $t(4)$ random variable takes value less than about -1.22, which is about 14.5%.

2. CUHK and HKU students took part in a math exam and obtained the following scores:

team	number of students	mean score	standard deviation
CUHK	80	84	12.4
HKU	100	80	11.2

   Assuming the students were chosen at random calculate the p-value for the hypothesis "CUHK does better than HKU in math" if the listed standard deviations are (a) the actual ones and (b) the adjusted sample deviations.

   Solution:

   (a) Let μ_A, μ_B be the means of the CUHK and HKU students, respectively, and σ_A, σ_B be their actual standard deviations. Then $\overline{A} - \overline{B}$ is a Normal$(\mu_A - \mu_B, \sqrt{\sigma_A^2/80 + \sigma_B^2/100})$ random variable. Under the null hypothesis that $\mu_A = \mu_B$, $\overline{A} - \overline{B}$ is a Normal(0, 1.782) random variable. The p-value for the hypothesis is $\mathrm{P}(\text{Normal}(0, 1.782) > 84 - 80) \approx \mathrm{P}(\text{Normal}(0, 1) > 2.244) \approx 0.012$, so we have about 98.8% confidence in the alternative hypothesis.

   (b) Let S_A, S_B be the adjusted sample deviations. Now the random variable

$$T = \frac{\overline{A} - \overline{B}}{\sqrt{\frac{1}{m} + \frac{1}{n}}\sqrt{\frac{m-1}{m+n-2}S_A^2 + \frac{n-1}{m+n-2}S_B^2}}$$

   is of type $t(m + n - 2)$, where $m = 80$ and $n = 100$ are the number of samples from the two populations. Plugging in the numbers the p-value is $\mathrm{P}(t(178) > 2.270) \approx 0.012$. The sample size here is sufficiently large that the two answers are the same.

[1] It is also reasonable to set $\mu \geq 26$ as the null hypothesis. The conclusion will be the same although the reasoning is a bit more involved.

3. A cookie manufacturer wants to test if replacing milk chocolate with dark chocolate in their product will lower the calorie count. To do so it creates sixteen cookie batches, tests the samples, and obtains the following numbers:

with dark chocolate	113	120	138	120	100	118	138	123
with milk chocolate	138	116	125	136	110	132	130	110

(a) If the batches were produced by eight (independent) cooks, each of which made one with dark and one with milk chocolate, which test would be appropriate to use, and what is the p-value?

Solution: As each pair of samples is not independent the paired t-test is more appropriate. The differences $D_i = MILK_i - DARK_i$ in calorie count are

$$25, -4, -13, 16, 10, 14, -8, -13.$$

We model these as independent normals of same (unknown) mean and variance. The random variable $\overline{D}/(S/\sqrt{8})$ is then a $t(7)$ random variable. As $\overline{D} = 3.375$, $S \approx 14.657$, and $\overline{D}/(S/\sqrt{8}) \approx 0.651$ we get a p-value of $\mathrm{P}(t(7) \geq 0.651) \approx 0.27$.

(b) If one cook produced all the dark chocolate batches and another one produced all the milk chocolate ones, which test would you use? How does the p-value change?

Solution: We can now use the independent sample test (as in question 2(b)). For the dark chocolates, the sample mean is 121.25 and the adjusted sample standard deviation is 12.52. For the milk chocolates, the sample mean is 124.63 and the adjusted sample standard deviation is 11.30. The p-value is given by $\mathrm{P}(t(14) \geq 0.567) \approx 0.29$, so we are now slightly less confident in the conclusion.

4. You are given two samples X_1, X_2 of a Uniform$(0, \theta_1)$ random variable and two samples Y_1, Y_2 of a Uniform$(0, \theta_2)$ random variable, all independent. You want to test the alternative hypothesis $\theta_2 > \theta_1$ against the null hypothesis $\theta_2 = \theta_1$. Consider the test T that accepts if $\min\{Y_1, Y_2\} > \max\{X_1, X_2\}$ and rejects if not.

(a) What is the false positive error of T?

Solution: False positive error occurs when $\min\{Y_1, Y_2\} > \max\{X_1, X_2\}$ and $\theta_1 = \theta_2$. As all 4! orderings of X_1, X_2, Y_1, Y_2 are equally likely and a false positive occurs only in $2! \cdot 2!$ of them (the Ys must come after the Xs but each group can be ordered arbitrarily), the false positive probability is $(2! \cdot 2!)/4! = 1/6$.

Alternatively, we can obtain the same answer by the total probability theorem. Let $M = \max\{X_1, X_2\}$. Then the conditional probability $\mathrm{P}(\min\{Y_1, Y_2\} > M | M = m)$ equals $\mathrm{P}(\min\{Y_1, Y_2\} > t) = (1 - t/\theta_1)^2$. The PDF of M is

$$f_M(m) = \frac{d}{dm} \mathrm{P}(X_1 \leq m, X_2 \leq m) = \frac{d}{dt} \mathrm{P}(X_1 \leq m)\, \mathrm{P}(X_2 \leq m) = \frac{d}{dm}(m/\theta_1)^2 = 2m/\theta_1^2.$$

By the total probability theorem,

$$\begin{aligned} \mathrm{P}(\min\{Y_1, Y_2\} > \max\{X_1, X_2\}) &= \int_0^{\theta_1} f_M(m) \cdot \mathrm{P}(\min\{Y_1, Y_2\} > M | M = m) dm \\ &= \int_0^{\theta_1} 2m/\theta_1^2 \cdot (1 - m/\theta_1)^2 dm \\ &= \frac{1}{6}. \end{aligned}$$

(b) What is the power (acceptance probability) of T as a function of $\rho = \theta_2/\theta_1$?

Solution: The test accepts when $\min\{Y_1, Y_2\} > \max\{X_1, X_2\}$. We apply the total probability theorem conditioning on the number N of samples Y_1, Y_2 that fall in the range $[0, \theta_1]$:

$$\begin{aligned}
\mathrm{P}(+) &= \mathrm{P}(+|N=0)\,\mathrm{P}(N=0) + \mathrm{P}(+|N=1)\,\mathrm{P}(N=1) + \mathrm{P}(+|N=2)\,\mathrm{P}(N=2) \\
&= 1 \cdot (1-\rho^{-1})^2 + \frac{1}{3} \cdot 2\rho^{-1}(1-\rho^{-1}) + \frac{1}{6} \cdot \rho^{-2} \\
&= 1 - \frac{4}{3\rho} + \frac{1}{2\rho^2}.
\end{aligned}$$

Justification: Clearly the test always accepts $Y_1, Y_2 > \theta_1$. When exactly one of the Y_i is in $[0, \theta_1]$ then a the test accepts when Y_i comes out last in the ordering of X_1, X_2, Y_i, an event of probability $1/3$. When X_1, X_2, Y_1, Y_2 are all in $[0, \theta_1]$ then for the test to accept Y_1 and Y_2 must come after X_1 and X_2 which happens with probability $1/6$.

Alternatively we can derive the same result using the total probability theorem for continuous random variables. By the total probability theorem

$$\begin{aligned}
\mathrm{P}(\min\{Y_1, Y_2\} > \max\{X_1, X_2\}) &= \int_0^{\theta_1} f_M(m) \cdot \mathrm{P}(\min\{Y_1, Y_2\} > M | M = m) dm \\
&= \int_0^{\theta_1} 2m/\theta_1^2 \cdot (1 - m/\theta_2)^2 dm \\
&= 1 - \frac{4\theta_1}{3\theta_2} + \frac{\theta_1^2}{2\theta_2^2}.
\end{aligned}$$

(c) **(Optional)** What is the likelihood ratio test, i.e., the test of the form

$$\frac{\sup_{H_1} f(x_1, x_2, y_1, y_2 | \theta_1, \theta_2)}{\sup_{H_0} f(x_1, x_2, y_1, y_2 | \theta_1, \theta_2)} \geq t$$

(where f is the joint PDF of the four samples) for a $1/8$ false positive error?

Solution: Under hypothesis H_0, θ_1 equals θ_2 and the joint PDF is maximized when $\theta_1 = \theta_2 = m$ where $m = \max\{x_1, x_2, y_1, y_2\}$, and $\sup_{H_0} f(x_1, x_2, y_1, y_2 | \theta_1, \theta_2) = 1/m^4$. Under H_1, θ_2 must be greater than θ_1 and we consider two cases. Let $m_x = \max\{x_1, x_2\}$ and $m_y = \max\{y_1, y_2\}$. When $m_x > m_y$ the supremum is attained at $\theta_1 = \theta_2 = m$, and it equals $1/m^4$. When $m_y \geq m_x$ it is attained at $\theta_1 = m_x$ and $\theta_2 = m_y$, and it equals $1/m_x^2 m_y^2$. Combining all this we obtain

$$\frac{\sup_{H_1} f(x_1, x_2, y_1, y_2 | \theta_1, \theta_2)}{\sup_{H_0} f(x_1, x_2, y_1, y_2 | \theta_1, \theta_2)} = \begin{cases} 1 & \text{if } m_x > m_y, \\ m_y^2/m_x^2, & \text{if } m_x \leq m_y. \end{cases}$$

The right hand side is always at least one, so the likelihood ratio test should accept when $\max\{Y_1, Y_2\}/\max\{X_1, X_2\}$ exceeds some threshold $t \geq 1$ and reject otherwise. It can be calculated that under the null hypothesis, $\mathrm{P}(\max\{Y_1, Y_2\} > t \max\{X_1, X_2\}) = 1/2t^2$ (assuming $t \geq 1$). For a false positive error of $1/8$, t should be set to 2.