1. A population of Douglas fir is normally distributed with respect to height, with μ = 40.0 m and σ = 10.0 m.

a) If you select a tree at random from this population, what is the probability that its height will be greater than 40.0 ***or*** less than 20.0?

$$z = \frac{X - \mu}{\sigma}$$

X = 40.0 ≅ z = 0 $P(x > 40.0) = P(z > 0) = 0.500$

X = 20.0 ≅ z = -2 $P(x < 20.0) = P(z < -2) = 0.0228$

$$P(x > 40.0 \text{ or } x < 20.0) = P(z > 0) + P(z < -2)$$
$$= 0.500 + 0.0228 = 0.523$$

b) If you select two more trees at random, what is the probability that the height of the first will be greater than 40.0 cm ***and*** the height of the second will be less than 20.0 cm?

$$P(x > 40.0 \text{ and } x < 20.0) = P(z > 0) \times P(z < -2)$$
$$= (0.500) \times (0.0228) = 0.0114$$

2. Research question: "Does stream restoration affect growth rates of juvenile coho salmon?" You decide to address this question by measuring growth rates in years before vs. after stream restoration activities. Your null hypothesis is that mean growth rates before vs. after restoration are equal, H_0: $\mu_1 = \mu_2$. Explain in terms of the research question the meaning of $\alpha = 0.05$.

At $\alpha = 0.05$ there is a 5% probability of concluding that stream restoration affects juvenile coho grow rates when restoration actually has no effect.

3. For samples of 10 paired measurements, the mean difference ($\bar{d}$) is 21.0, and the variance (s^2) is 250.00.
What are the 95% confidence limits for the mean difference (μ_d)?

95% CL: $\bar{d} \pm t_{0.05,\nu} s_{\bar{d}}$ ν=n−1 = 9 $s_{\bar{d}} = \sqrt{s^2/n} = \sqrt{250.00/10} = 5.0$

= 21.0 ±(2.262)(5.0)
= 21.0 ±11.3
[9.7, 32.3]

4. Research question: was summer 2008 significantly wetter than summer 2007? Cumulative precipitation during the months of June-Sept. was recorded in 20 western Washington cities during both summers. Examination of independent data suggests that the precipitation levels were normally distributed.
a) Which statistical test is most appropriate? Paired sample t-test, one-tailed
b) Why?
Two samples, observations paired by location,
One-tailed because query regards only whether one summer was drier than the other.
Two-tailed test would address whether summers *differed* (drier or wetter) in precipitation.
Normal distribution => parametric test more powerful than nonparametric

c) State suitable null and alternative hypotheses that could be used to address the research question.

$H_0 : \mu_d \le 0$ $H_A : \mu_d > 0$, where $d_i = X_{i,2008} - X_{i,2007}$

5. Research question: do bigleaf maples produce leaves in high light environments that differ in size from leaves produced in low light? You recorded the area of ten randomly selected sun leaves and 16 randomly selected shade leaves. Independently collected data suggest that leaf areas were normally distributed, with equal variances among leaves in high and low light environments.

$H_0: \mu_1 = \mu_2$	$\bar{X}_1 = 211.0$	(1=shade)
$H_A: \mu_1 \neq \mu_2$	$\bar{X}_2 = 208.0$	(2=sun)
$\alpha = 0.05$	$s_{\bar{X}_1 - \bar{X}_2} = 1.2$	

Complete a statistical analysis to answer the research question. For full credit, include the following:
1) Calculate the test statistic, t_{calc}. Show all formulas and values used.
2) State your statistical decision about the null hypothesis.
3) State your conclusion, in words, about the research question.

1)

$$t_{calc} = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{211.0 - 208.0}{1.2} = \frac{3.0}{1.2} = 2.5$$

2) $\nu = n_1 + n_2 - 2 = 16 + 10 - 2 = 24$

$t_{0.02(2),24} = 2.064$

$t_{calc} > t_{0.02(2),24}$ $P < \alpha$

Reject H_o; accept H_A, $\mu_1 \neq \mu_2$

3) Bigleaf maple leaves grown in low light (μ_1) are larger than leaves grown in high light (μ_2).

6. Suppose independently collected data suggest that the distribution of leaf areas in problem 4 deviates severely from normality.

a) Which test would provide greatest power to answer the research question?

Mann-Whitney two-sample test.
Use nonparametric test that does not assume normality.
Mann-Whitney test because two-sample study design; no mechanism of pairing evident.

b) State the hypotheses appropriate for that test.

H_o; bigleaf maple leaves grown in high light environments are equal in size to bigleaf maple leaves grown in low light environments.

H_A: bigleaf maple leaves grown in high light environments differ in size with bigleaf maple leaves grown in low light environments.

7. During the m&m example in class, I instructed one of your peers to select ten plain m&ms randomly relative to their positions. Describe how you could have selected those ten m&ms in a way that would have been random relative to position.

Many procedures would ensure randomness (relative to position). To qualify as “random,” each m&m must have an equal probability of being included in the sample, regardless of position.

Possibilities:

(1) Assign each m&ms a number, and then select a random sample of those numbers – e.g., with a random number generator (w/ uniform distribution).

(2) Select all m&ms of a given color (if there are exactly ten of a particular color). Because color and position should be independent, this sample should be random relative to position.

(3) …