BUDT 730 DMD
Practice Final Exam

Question 1

a) What is the sampling distribution of the average income from the parking lot over 45 days?

Normal ($850, $150/$\sqrt{45}$)

b) Write out the null and the alternative hypotheses for the manager. Clearly state them in terms of your parameter.

Let μ = the true daily mean income
H_o: $\mu \geq \$850$
H_a: $\mu < \$850$

Example of wrong answers:
H_0: The attendant is not cheating, $\mu \geq \$800, \$38{,}250, \$36{,}000, \leq 2{,}250$, ...
H_a: The attendant is cheating, $\mu > \$850$, ...

c) What is the risk of Type I and Type II error? Clearly state them in English.

Type I: The mean income is $850 (the attendant is not cheating), but we conclude that it is less than $850 (the attendant is cheating) and fire the attendant.

Type II: The mean income is less than $850 (the attendant is cheating), but we conclude that it is equal to $850 (the attendant is not cheating) and keep the attendant.

Which one is worse? Type II
Type I: The loss is $1,000
Type II: The loss is $50/day. If we keep him more than 20 days, the loss will be greater than $1,000

How does this influence our significance level we need?
Set the type II error low -> use high α (0.1 or higher)

d) ASSUME that the p-value is 0.0127. Write out the meaning of this p-value in English.

Assuming that the true daily mean income is $850, there is a 1.27% chance that on average the attendant turns in $800/day or less, over the course of 45 days.

Example of wrong answers:
Assuming that the true mean is $850, there is 1.27% chance that on average we get $800 or higher. (It should be an average over 45 days)
Assuming that H_0 is true, there is 1.27% chance that on average we get the result this extreme or more. (You need to specify the extreme value)

(e) Set up a test statistic for the hypothesis test and compute the value

Since the true standard decision is known, we use Z statistics:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{45}}} = \frac{800 - 850}{\frac{150}{\sqrt{45}}} = -2.2$$

(f) Using a 5% level of significance, should the manager fire the attendant? (Use the critical value method)

$$|-\mathbf{2.2361}| > \boldsymbol{Z_{0.95}} = \mathbf{1.645}$$

We reject the null and fire the attendant

Question 2

a) Given the variables specified above, name two variables for which it would be reasonable to convert to dummy variables to be used in a regression model.

Age, History, or Catalogs; Gender and Close are already dummy variables in their current forms.

b) Is this model valid as an explanatory model? Why or why not. Explain your answer.

The model is not valis. The residual plot shows that the residual is non-normal with non-constant variance.

b) Provide an interpretation of the p-value for the Salary.

The small p-value implies that the coefficient of Salary variable is not equal to zero. Therefore, we should include Salary in our regression model.

c) Specify two pieces of information from these results that support the belief that this model is an improvement over the initial model?

1. The multiple-R and the R-square values are improved.
2. All p-values are small.
3. Nonlinear patterns in fitted vs. actual Log(AmountSpent) have disappeared and residuals vs. fitted values shows that the variance of residual is now constant (residual is homoscedastic).

Note: You cannot use the standard error to compare the two models.

d) Provide an economic interpretation for the coefficient of the Catalogs = 18 variable.

On average the customers who received 18 catalogs spend 57% more compared to the customers who received 6 catalogs, when the rest of conditions remains the same

e) Predict AmountSpent for a customer with a household income of $100,000 who lives close to stores that sell similar merchandise, was a high spender in the previous year, and received 6 catalogs. You may round each coefficient to the hundredths place. You must show your work to receive full credit.

f)

Log(AmountSpent)=5.9962 -0.2741 + 1.3796+0.0871=7.1888

AmountSpent = $e^{7.1888}$

Question 3

(a) What is the price elasticity of demand (the change in demand in relation to a change in its price) on routes on which Southwest is present?

Price Elasticity = Coef of log(FARE) + Coef of SW*log(FARE) = -0.2703-0.3535= -0.6238

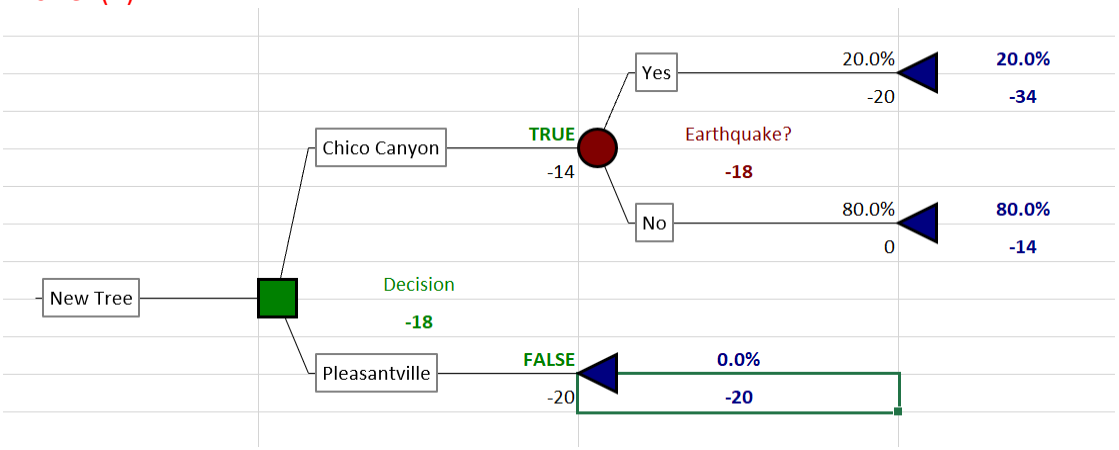
(b) Predict demand for a route on which Southwest does not fly and the fare is $500. You may round each coefficient to the hundredths place. You must show your work to receive full credit.

Log (PAX(demand)) = 10.52 +-0.27*ln(500) = 8.8421

PAX = exp(8.8421)= 6919

Question 4

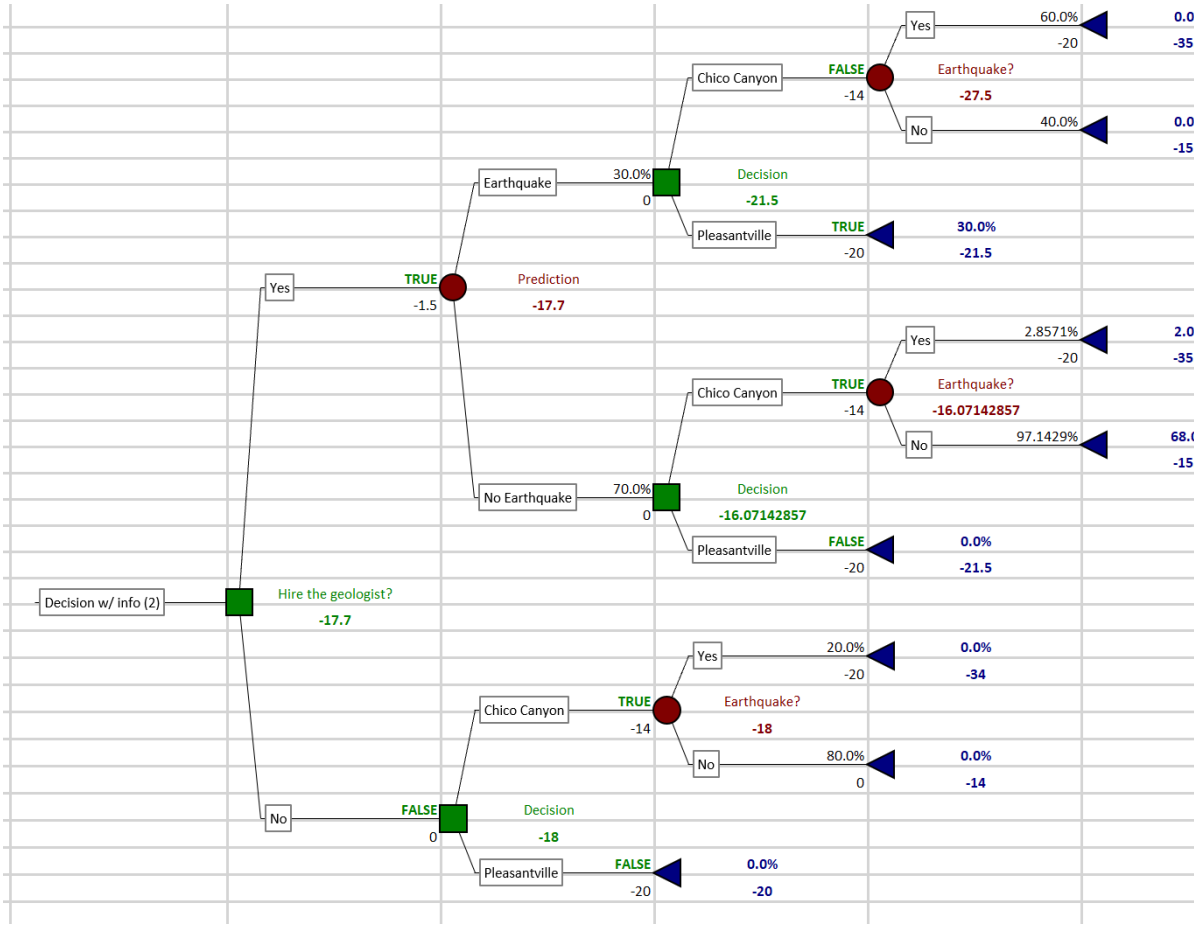
Answer (A)



EMV at the chance node: -14+[(-20)*0.2+0]=-14-4=-18
EMV at the decision node: max(-18, -20)=-18

(B) The plant should be built in Chico Canyon. Even with the earthquake uncertainty, the expected cost is still lower ($18m vs. $20m).

(C)



(D)

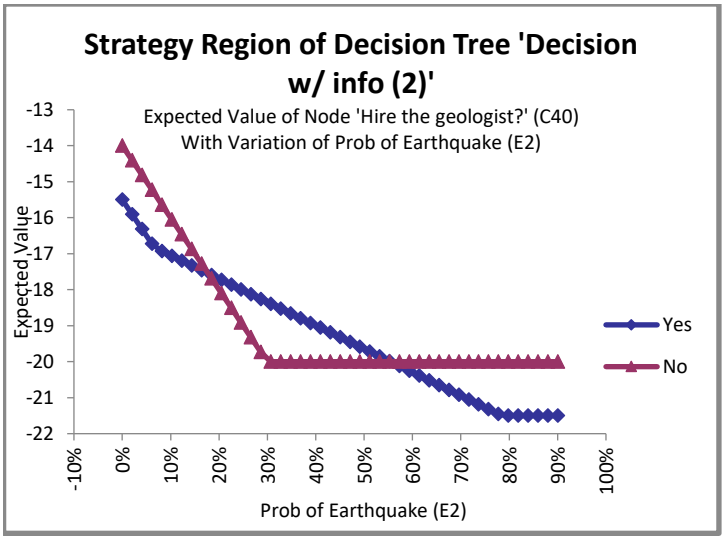
Where should the power company build the plant?

The decision depends on the first stage decision and the geologist's prediction. Suppose that the geologist is hired. The plant should be built at Pleasantville if he predicts an earthquake at Chico. Otherwise, it should be built at Chico. If the geologist is not hired, it should be built at Chico.

Should the company hire the geologist? What is the EVI?

The tree shows that the geologist's information is very valuable. The information lowers the expected cost to $17.7m from $18m. The EVI is (-17.7+1.5) – (-$18 m) = $1.8 mil. Therefore, the company would be justified in paying the geologist's fee.

(E)



The first stage decision (hiring geologist) changes from no to yes at p (probability of earthquake) =15% and yes to no at p=56%.

The second stage decision
For ' Yes' decision at the first stage:

p< 5%: Build the plant in Chico Canyon regardless of the prediction outcome.
5%< p <78%: Depending on the prediction outcome.
p > 78%: Build the plant in Pleasantville regardless of the prediction outcome.

For ' No' decision at the first stage:
p <30%: Build the plant in Chico Canyon.
p > 30%: Build the plant in Pleasantville.