question **37** views

question on problem set 6

I'm not sure what I've done wrong for problem 1a in problem set 6. Here is the R code I used.
It seems that the answer according to the way I did this falls between 2000 and 2001 (entering between 5-6 as the year in the Pred function). Does anyone know what I did wrong? This problem set question only gives you 1 chance, not multiple as the labs sometimes do, so I just want to understand what I did incorrectly, since I already lost the credit.

Thanks.

```
> brazil <- world[world$Country == "Brazil",]
> brazil_select <- brazil[brazil$year >= 1995 , ]
> brazil_select$mobile.mil <- brazil_select$mobile.users / 1000000
> brazil_select$time <- brazil_select$year - 1995
> expFit(brazil_select$time, brazil_select$mobile.mil)
 a =  3.22757
 b =  1.33871
 R-squared =  0.93243
> logisticFit(brazil_select$time, brazil_select$mobile.mil)
Logistic Fit
 C =  347.9316
 a =  72.74891
 b =  1.3595
 R-squared =  0.99785
> logisticFitPred(brazil_select$time, brazil_select$mobile.mil, 5) -- gave the 20.88 -- chose 5 since 2000 is 5 years since 1995
--given the correct answer is 23.19, I tried some other years close by to what I thought I should enter, to see if it would give the 23.19.

> logisticFitPred(brazil_select$time, brazil_select$mobile.mil, 10) -- gave 79.56
> logisticFitPred(brazil_select$time, brazil_select$mobile.mil, 7) -- gave 36.72
> logisticFitPred(brazil_select$time, brazil_select$mobile.mil, 6) -- gave 27.78
```

1a. Find the number of mobile users in Brazil (in millions) in 2000, using R. *(Round to 2 decimal places.)*
20.88 - incorrect
Ans. 23.19

problem_set6

4 days ago by Karen West

the students' answer, *where students collectively construct a single answer*

You don't have to predict the number of users in 2000. The answer is found in your data.

4 days ago by SeaBreeze

followup discussions *for lingering questions and comments*

Resolved

Unresolved

Karen West 2 days ago

Using this R code for Problem Set 6, #1e, I was off by 3 - does anyone know what I did wrong?

```
> time <- brazil_select$time - 1995
> muMil <- brazil_select$mobile.mil
> logisticFitPred(time,muMil,30)
```

1e. Using the best-fitting model, predict the number of mobile users (in millions) in Brazil in 2025. *(Round to zero decimal places.)*
348 - incorrect

348

345

Karen West 2 days ago The output of the R code above was the predicted value in the logistic function of: 347.932 or 348.

Anonymous 2 days ago Karen, the 3 lines of code look OK to me. So I strongly suspect that something is wrong with one of your vectors time or muMil. Inspect them. If you're not sure, post the values and how they were created.

Karen West 1 day ago Here are the 2 vectors I used in the R code above - it appears that the time vector is messed up?

```
> muMil
 [1]   1.285533   2.498154   4.550000   7.368218  15.032698  23.188171  28.745769  34.880964  46.373266  65.605000
[11]  86.210336  99.918621 120.980103 150.641403 169.385584 196.929978 234.357507 248.323703
> time
 [1] -1995 -1994 -1993 -1992 -1991 -1990 -1989 -1988 -1987 -1986 -1985 -1984 -1983 -1982 -1981 -1980 -1979 -1978
>
```

I had already subtracted 1995?

```
> brazil_select$time
 [1]  0  1  2  3  4  5  6  7  8  9 10 11 12 13
[15] 14 15 16 17
>
```

Anonymous 1 day ago Karen, looks like you subtracted 1995 twice!

Resolved

Unresolved

Karen West 23 hours ago

Problem Set 6, Question 2c: I got a slightly different answer using R verses by hand and I want to know what I'm doing wrong with my math? Does anyone know?

2c. Predict the number of cases of flu on **Day 14** (when "Day" is equal to 14), using the exponential model. *(Round to zero decimal places.)*

Typing the data set into R:

> expFit(fcDay,fcFlu.Cases)

a = 76.63569

b = 1.46096

R-squared = 0.98409

> expFitPred(fcDay,fcFlu.Cases, 14)

predicted 15466 flu cases.

how do you do this without R? I got the answer incorrect by:

exponential function: f(t) = ab^t
*f(t) = 76.64 * (1.46)^t*
*f(14) = 76.64 * (1.46)^14*
which would be 15324 flu cases on day 14.

Anonymous 21 hours ago forget the exact solution. The grader expects the rounded value of b to be used. You didn't round properly in your last step.

Karen West 16 hours ago So the correct answer I should submit is my hand calculated value, the 15324, rather than the value R calculated, which is not using rounded values?
You only get one chance, and I thought, do I use the value that R calculated, or the value I calculated by hand? It sounds from your response that I should use the rounded value of b that I used in the hand calculation.

Karen West 16 hours ago yet another problem incorrect! It can be frustrating when you are 1 off and you attempted to inquire first to make sure with only 1 try to get it right!
Oh well -- I was one off with my hand calculation and I do not know why.

2c. Predict the number of cases of flu on **Day 14** (when "Day" is equal to 14), using the exponential model. *(Round to zero decimal places.)*
15324 - incorrect

15324

15325

Anonymous 15 hours ago Karen, as I said before, you did not round properly in your last step. Any calculator gives you 15324.82

Karen West 15 hours ago I did not actually know that you meant the hand calculation was rounded incorrectly. I thought you had meant the R calculation,
which does not round a and b. So it did not occur to me to go back to check the hand calculated value. I did that after I got it incorrect and yes it is exactly as you said - any calculator would tell you that I rounded incorrectly in my hand calculation.
I guess I interpreted your comment incorrectly, in that I thought you were comparing saying R did not round and my hand calculation did - the reason for the discrepancy in values.

And I've only been interrupted 50 billion times today - the kids are off from school both today and tomorrow due to 20 inches (so far!) of snow. Other parts of MA got 33-34 inches - so lots of snow blowing and shovelling around here today! No excuse of course for carelessness - I often make careless mistakes when tired. ;-)

Thank you for responding! I'm so far behind that I may never finish on time, but I'm trying to get there. Otherwise I will have wasted

a lot of time trying but not making the deadline on time. ;-)

Anonymous 1 hour ago you're welcome! hopefully you'll get 70% or better! Hang in there.
hopefully the discussion forum will stay open, after the deadline, for more questions so you can finish. Coursera allowed that for the DASI course.

Karen West 48 minutes ago What is a DASI course on Coursera?
Does EdX work the same in this way?
I hope to finish by Feb. 11th, but at the rate I've been going recently, not so sure, but I'll try. ;-)

Anonymous 33 minutes ago see my review of the DASI course https://piazza.com/class/i0qpzfmi1eq6ca?cid=857 New class starts in March.
the instructor is a statistician, and is more rigorous with the language, which helps everyone. Don't have to worry much about rounding, etc.
I think people would agree that it is a harder course, but not much harder, for the right reasons fortunately, but it is still introductory. There is no better introductory stats MOOC than DASI. Much better participation in the discussion forum by the instructor and Community TAs (CTAs). LABs consist of gentle R commands.
There is MUCH overlap with this course, but spends more time on probability. Thus you should be able to pass the DASI course.

Karen West 1 minute ago Thanks - I'll keep that in mind for future reference if I want to study any further in this area! ;-)

Resolved

Unresolved

Karen West 1 hour ago
Apologies in advance for: the need for math help, and also, for being behind - still trying to catch up! ;-)

New Question on Problem Set 6 question 3c:

3c. Another researcher assumed that the wolves would experience exponential growth because there were no predators. He fit an exponential model to this data. What is the growth factor for his model? *(Round to 2 decimal places.)*
1.8 - incorrect

1.8

1.34
Can someone correct my math and exponential model calculation here?

I knew that the exponential model was f(t) = a*b^t
I tried at first to calculate the percent change from year 1 to year 3 in wolf population, 20/25 = 0.8 = r, with r being the proportional change of the outcome variable for every whole unit of increase on the input variable.
With b = 1 + r, this is why I put the b growth factor as 1.8.
It was incorrect, so I thought, perhaps it's because I should have only gone from year 1 to year 2, one unit of change in year, even though that is not in the table.
So that would give a change of 10/25 = 0.4 making b = 1.4 - but you cannot assume in the exponential model that at year 2, you will have 10 more wolves as you could with the linear model.
So I then tried to calculate the values of a and b for the exponential model.
f(0) = a*b^0 = a = 15 wolves at year 0
f(3) = 45 = 15 * b^3
3 = b^3
log(3) = 3 * log(b)
0.1590404 = log(b)
10^0.1590404 = b
b = 1.4422
So then I plugged these values of a and b into f(t) = a*b^t for t = 2, one year unit of change:
f(2) = 15 * (1.4422)^2 = 31.1991126 = 31.2
years 1->2: 31.2 - 25 = 6.2 increase in wolves with exponential model year 1 to 2
so percent change: 6.2/25 = 0.248
That is not the correct answer either, since that would make b=1.248, and the answer is 1.34.
If someone could help clear up this confusion, thank you in advance! ;-)

Anonymous 56 minutes ago far simpler to let R do the work for you:

```
x <- c(1,3)
y <- c(25,45)
(wolvesexpfit <- expFit(x, y))
```

but discretely:
f(1) = 25, f(3) = 45
a*b^1 = 25
a*b^3 = 45
45/25 = (a*b^3)/(a*b^1) = b^2 ---> b = 1.xx

Resolved

Unresolved

Karen West 18 minutes ago
New question on 3g: This time I used R?

3g. Using the best-fitting model, how many years must pass before there are **more than** 325 wolves in Yellowstone? *(Round to zero decimal places.)*
84 - incorrect

84

10
I thought the best fitting model was exponential, since R^2 = 1?
And the problem I got correct, 3f said:
3f. By 2002, there were 147 wolves in Yellowstone Park. Which model was determined to fit the data better?
Exponential Exponential - correct Linear Neither model appears to fit well
So back to what I did for 3g in R:

```
> x = c(1,3,7)
> y = c(25,45,147)
> expFit(x,y)
 a =  18.58487
 b =  1.34361
 R-squared =  1
> expFitPred(y,x,326)
```

gave predicted year value: 83.895
Sorry to have to ask for help again on problem set 6, but if anyone can help, it's appreciated. I hope to start unit 7 today,
among other things! This is almost the end of unit 6. Thanks.

Anonymous 27 seconds ago expFitPred() should not be used for this problem. It can be used, but it is more work. Instead, simply solve for t in the exponential model that causes f(t)=325:
t <- log(wcnt/a)/log(b)
you get a decimal answer, but the question is kinda of tricky: it asks how many years needs to pass such that the wolf count exceeds 325. By solving for t in the exponential model equation, you will get a decimal answer that says 325 occurred DURING a particular year. However, the question really is asking, at the end of each year does the wolf count exceed 325? Well if you solved t=4.3, then it took 5 years for the threshold to be exceeded, because the wolf count is being sampled at the end of every year.