35 YEARS OF IMPAC(T³™)

T³IC 2023

Teachers Teaching with Technology™

International Conference

Starting Strong in AP® Statistics with the TI-84 Plus CE Graphing Calculator: Exploring Data and Random Variables

Daren Starnes
T^3 National Instructor
Hilton Head, SC

Josh Tabor
The Potter's School
Tucson, AZ

PDF of the slides available at tinyurl.com/t3online23jtdss

Agenda

- Exploring Data

  Fitting a linear model to data: what do r^2 and s tell us?
- Random Variables

  Calculating normal and binomial probabilities on the TI-83/84—What work do students need to show?

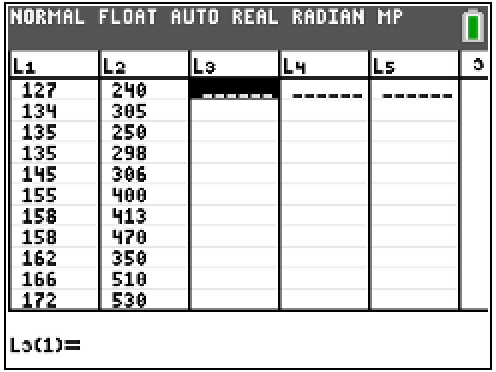
Exploring Data

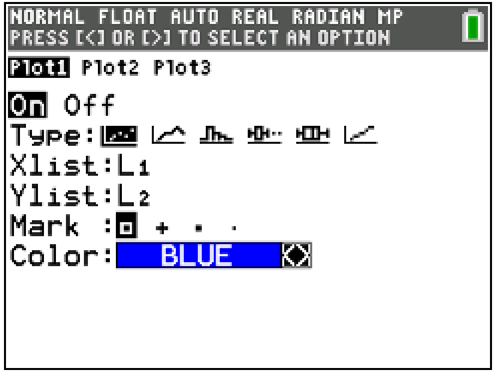
Bullfrogs and r^2

Based on #1 from the 2022 AP® Statistics exam

A biologist gathered data on the length, in millimeters (mm) and the mass, in grams (g), for 11 bullfrogs. Here are the data:

(a) Make a scatterplot to display the relationship between y = mass and x = length for these 11 bullfrogs. Describe what you see.



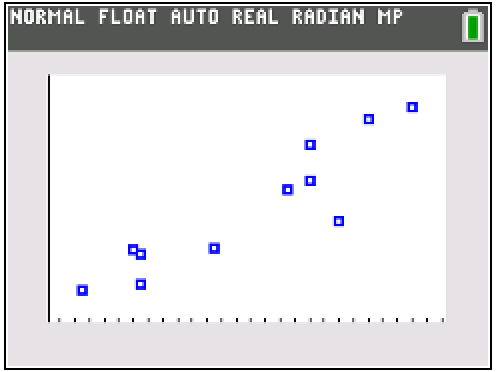


Length (mm)	Mass (g)
127	240
134	305
135	250
135	298
145	306
155	400
158	413
158	470
162	350
166	510
172	530

A biologist gathered data on the length, in millimeters (mm) and the mass, in grams (g), for 11 bullfrogs. Here are the data:

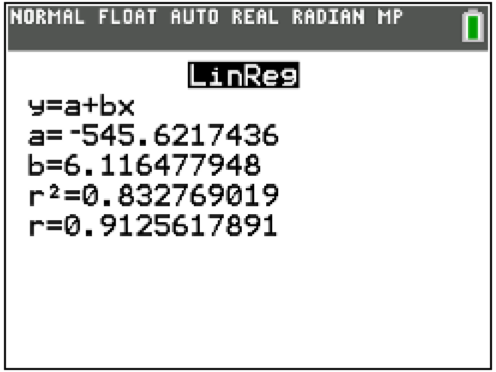
Length (mm)	Mass (g)
127	240
134	305
135	250
135	298
145	306
155	400
158	413
158	470
162	350
166	510
172	530

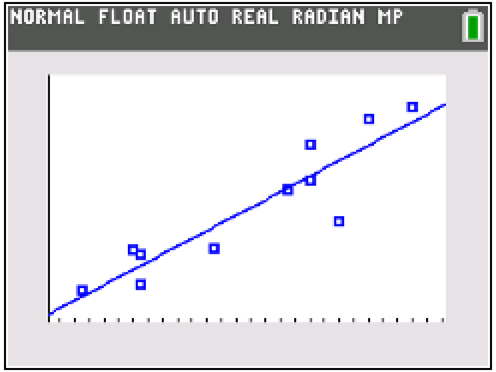
(a) Make a scatterplot to display the relationship between *y* = mass and *x* = length for these 11 bullfrogs. Describe what you see.



There is a moderately strong, positive, linear association between mass and length for these 11 bullfrogs, with a possible outlier at (162, 350).

(b) Calculate the equation of the least-squares regression line and show it on the scatterplot.





(c) How well does the line work?

There are two ways to answer this question:

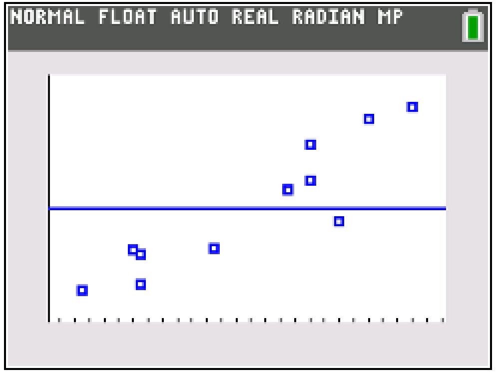
- With r^2, the coefficient of determination
- With s, the standard deviation of the residuals

But before we go any further, we need to be clear on *what we are comparing the line to.*

(c) How well does the line work?

Imagine we are going to randomly select a 12th bullfrog. What mass would we predict for this not-yet-selected bullfrog?

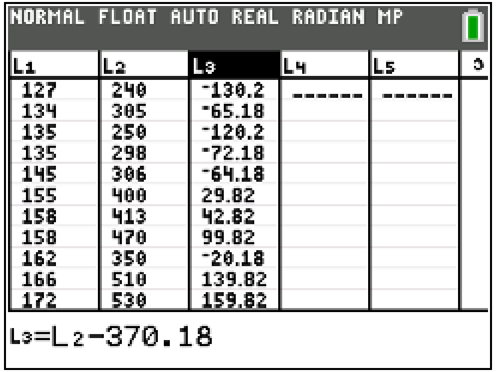
- Using one-var stats on L2, the mean mass, $\bar{y}$ = 370.18 g.

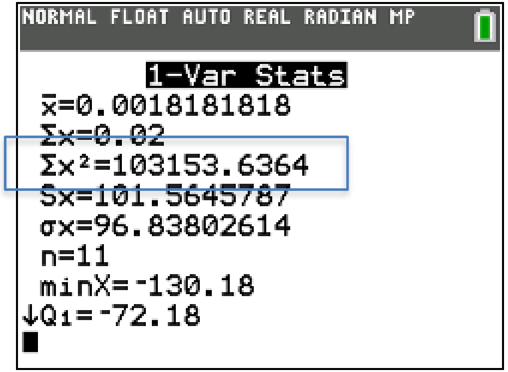


(c) How well does the line work?

Using this mean-only model, how good will our predictions be?

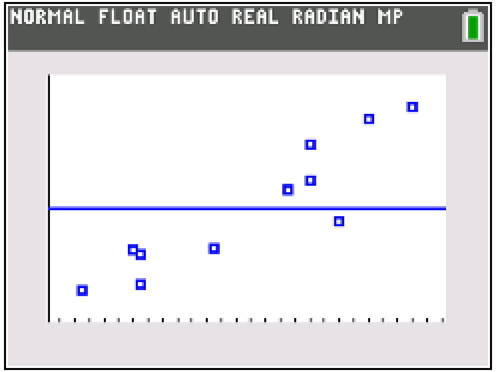
- In L3, find the residuals (deviations): L3 = L2 – 370.18.
- What is the average deviation?
- Using one-var stats on L3, average deviation ≈ 0! Why is that? How could we fix this problem?
- Use sum of squared residuals = 103,154

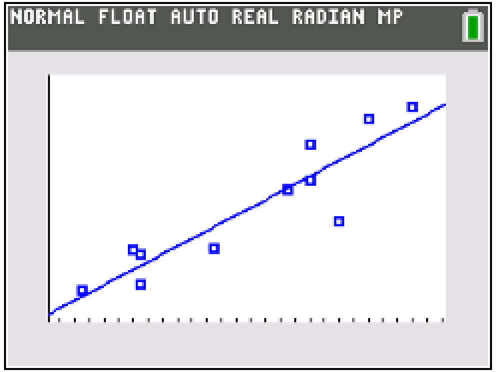




(c) How well does the line work?

What if we get to use the least-squares regression line to make predictions?

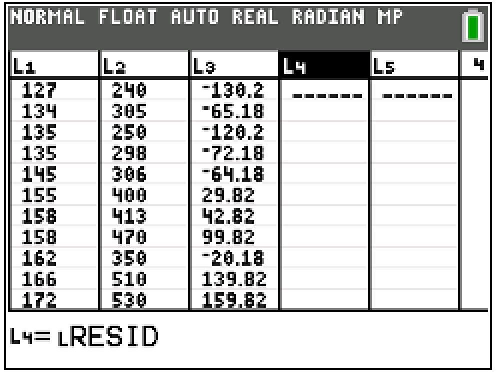




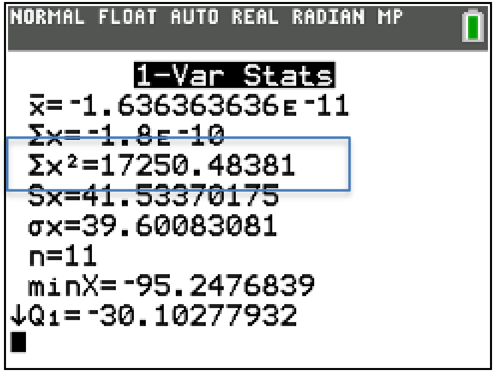
(c) How well does the line work?

What if we get to use the least-squares regression line to make predictions?

- In L4, place the residuals from the least-squares regression line using the RESID list.
- What is the average deviation? Again, ≈ 0.
- The sum of squared residuals = 17,250



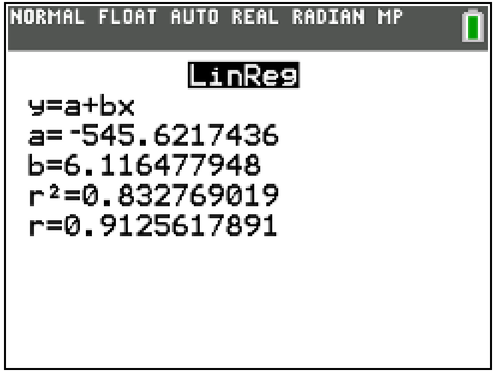
L1	L2	L3	L4	L5
127	240	-130.2	------	------
134	305	-65.18		
135	250	-120.2		
135	298	-72.18		
145	306	-64.18		
155	400	29.82		
158	413	42.82		
158	470	99.82		
162	350	-20.18		
166	510	139.82		
172	530	159.82		



(c) How well does the line work?

By how much did the sum of squared residuals go down when we used the least-squares regression line instead of the mean-only model?

- (103,154 – 17,250)/103,154 = 0.833 = r^2 (matches the output from LinReg)
- 83.3% of the variability in the mass of the bullfrogs is accounted for by the linear model using x = length.



(c) How well does the line work?

We can also use the sum of squared residuals from each model to calculate the relevant standard deviation:

$$s_y = \sqrt{\frac{(y-\bar{y})^2}{n-1}} = \sqrt{\frac{103{,}154}{11-1}} = 101.6\,g$$

$$s = \sqrt{\frac{(y-\hat{y})^2}{n-2}} = \sqrt{\frac{17{,}250}{11-2}} = 43.8\,g$$

- Using the least-squares regression line reduced the typical prediction error by more than half!

Random Variables

Doing the Floss: Grin and Bear It with Random Variables

Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

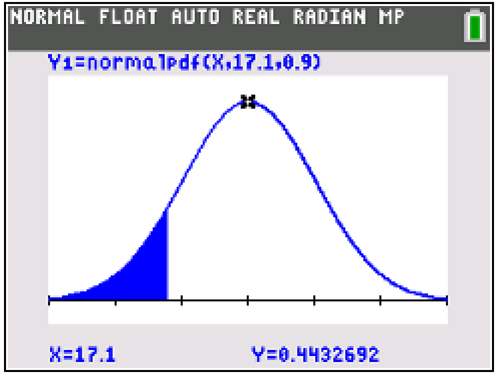
(a) From experience, Flossie needs at least 16 inches of floss to clean between all of her teeth. What is the probability that Flossie pulls off too little floss on a given day?

- Define the random variable.

Let X = the amount of floss (in inches) that Flossie pulls off on a given day

- State how the random variable is distributed.

X has a normal distribution with mean $\mu = 17.1$ inches and standard deviation $\sigma = 0.9$ inch



- Identify the values of interest.

We want to find $P(X < 16)$.

Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(a) From experience, Flossie needs at least 16 inches of floss to clean between all of her teeth. What is the probability that Flossie pulls off too little floss on a given day?

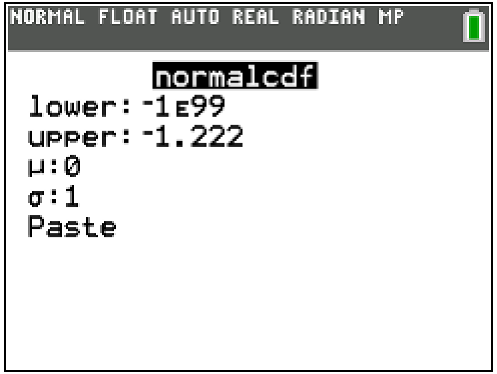
- Calculate the probability—show your work!

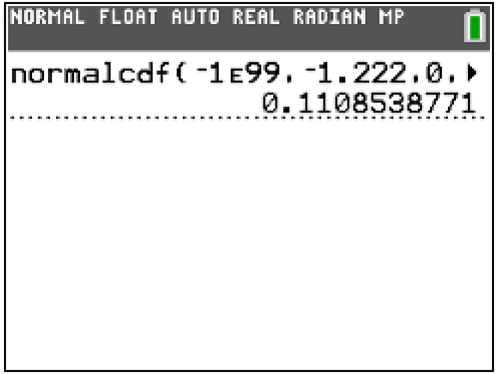
Using a formula

$$z = \frac{16-17.1}{0.9} = -1.222$$

TI-83/84: $P(Z < -1.222) = 0.1109$

Using Table A : $P(Z < -1.22) = 0.1112$



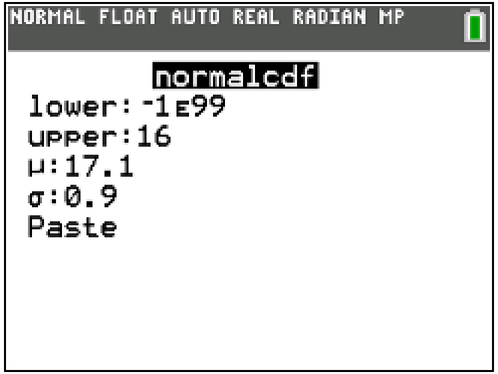


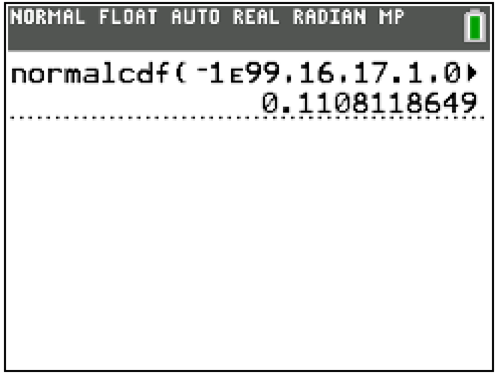
Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(a) From experience, Flossie needs at least 16 inches of floss to clean between all of her teeth. What is the probability that Flossie pulls off too little floss on a given day?

- Calculate the probability—show your work!

With technology $P(X < 16)$ = normalcdf(lower bound: 0, upper bound: 16, mean: 17.1, SD: 0.9) = 0.1108.





Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(b) The current month has 31 days. Find the probability that Flossie pulls off too little floss on more than 2 days in the month.

- Define the random variable.

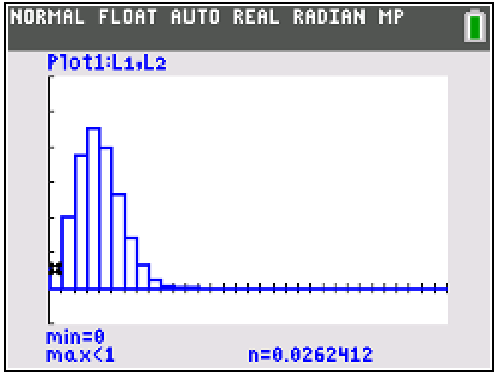
Let Y = the number of days in the current month when Flossie pulls off too little floss

- State how the random variable is distributed.

Y has a binomial distribution with n = 31 and p = 0.1108.

- Identify the values of interest.

We want to find $P(Y > 2)$.



Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(b) The current month has 31 days. Find the probability that Flossie pulls off too little floss on more than 2 days in the month.

- Calculate the probability—show your work!

Using a formula

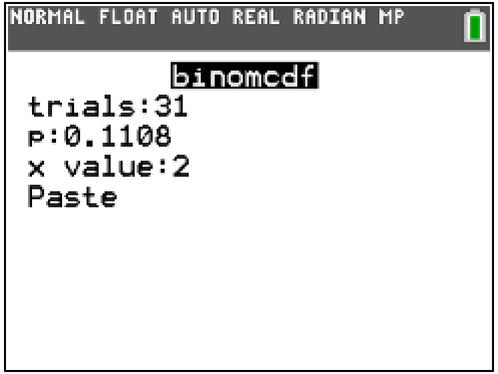
$$
\begin{aligned}
P(Y > 2) &= 1 - P(Y \le 2) = 1 - \left[P(Y=0) + P(Y=1) + P(Y=2)\right] \\
&= 1 - \left[\binom{31}{0}(0.1108)^0(0.8892)^{31} + \binom{31}{1}(0.1108)^1(0.8892)^{30} + \binom{31}{2}(0.1108)^2(0.8892)^{29}\right] \\
&= 1 - (0.0262 + 0.1014 + 0.1895) \\
&= 1 - 0.3171 \\
&= 0.6829
\end{aligned}
$$

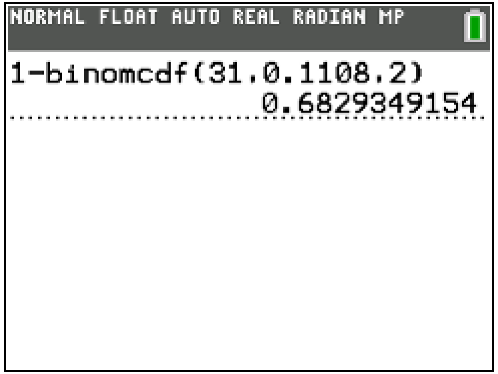
Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(b) The current month has 31 days. Find the probability that Flossie pulls off too little floss on more than 2 days in the month.

- Calculate the probability—show your work!

With technology $P(Y > 2) = 1 -$ binomcdf(trials: 31, p: 0.1108, x value: 2) = 0.6829





Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(c) At the beginning of the current month, the spool has 45 feet of floss. What is the probability that Flossie will run out of floss by the end of the month?

- Define the random variable.

Let $T = X_1 + X_2 + \ldots + X_{31}$.

- State how the random variable is distributed.

T is the sum of 31 independent and identically distributed normal random variables. So *T* is normally distributed with mean $\mu_T = \mu_{X_1} + \mu_{X_2} + \ldots + \mu_{X_{31}} = 17.1 + 17.1 + \ldots + 17.1 = 31(17.1) = 530.1 \text{ inches}$ and standard deviation $\sigma_T^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \ldots + \sigma_{X_{31}}^2 = 31(0.9^2) = 25.11 \Rightarrow \sigma_T = \sqrt{25.11} = 5.01 \text{ inches}$

- Identify the values of interest.

We want to find $P(T \geq 540)$ since 45 feet = 540 inches.

Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(c) At the beginning of the current month, the spool has 45 feet of floss. What is the probability that Flossie will run out of floss by the end of the month?

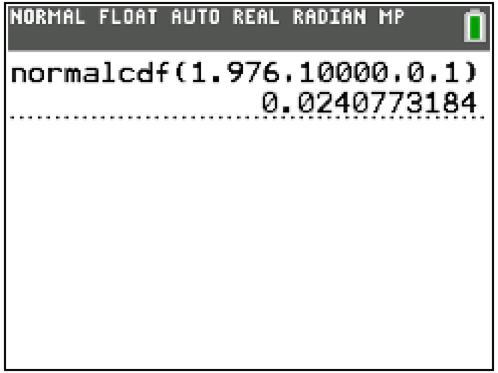
- Calculate the probability—show your work!

Using a formula

$$z = \frac{540 - 530.1}{5.01} = 1.976$$

TI-83/84: $P(Z \geq 1.976) = 0.0241$

Using Table A : $P(Z \geq 1.98) = 1 - 0.9761 = 0.0241$

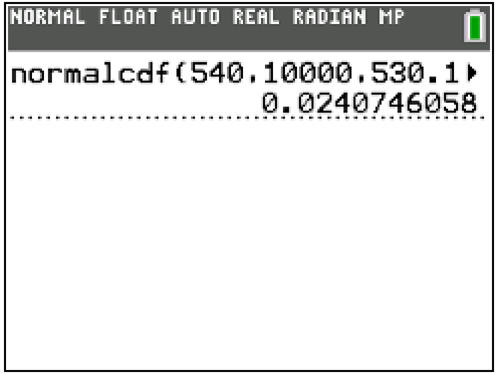


Flossie McFlossington flosses her teeth once per day. The amount of floss that Flossie pulls off the spool on a given day can be modeled by a normal distribution with mean 17.1 inches and standard deviation 0.9 inch.

(c) At the beginning of the current month, the spool has 45 feet of floss. What is the probability that Flossie will run out of floss by the end of the month?

- Calculate the probability—show your work!

With technology $P(T \geq 540)$ = normalcdf(lower bound: 540, upper bound: 1000, mean: 530.1, SD: 5.01) = 0.0241.



PDF of the slides available at tinyurl.com/t3online23jtdss