ESTIMATING CONTRACT TIMES FOR TRANSPORTATION PROJECTS: CREATING A STATISTICAL MODEL TO ESTIMATE TIMES USING BID QUANTITIES

CSCE, MAY 2017 GUILLERMO NEVETT PAUL GOODRUM UNIVERSITY OF COLORADO BOULDER

Presentation Agenda

Problem statement.
Current methods for time estimation
Explain the development process for the statistical model.
Path forward.

Problem Statement

Low accuracy



Approved FHWA Contract Time Determination Techniques

Bar Charts Method
 Critical Path Method
 Estimated Cost Methods

First Steps

 Clean and Organize Data to get same units and grouping similar variables (e.g. piping)
 Convert Engineer's Estimate to 2015 USD using the NHCCI index (https://www.fhwa.dot.gov/policyinformation/nhcci.cfm)

The data

K RStudio	File Edit	Code View Plots Session Build Debug Profile Tools	Window Help		, 🎅 🔽 🌒)) 100% 🛛	🕬 🔜 Wed 20:06 Q	. 💿 😑				
RStudio											
S Project: (None) - B Bill all / A Conto file/fun_1 + Addins -											
Deservenuel.R*+nen_19+newrund											
cont_id	sewer perfpir	Bid Item	Freq.	Bid Item	Freq.	2003 USD cont_id1	TOT_BID_AMT 2				
1401 C17846	5.8664680		177		00	8387268 C17846	882650!				
1402 C14606R	3.0445225	Crushed agaregate		Commercial Asphalt	20	8442246 C14606R	898118;				
1403 C18085	NA 6.8638	crostica aggiogato			20	8468312 C18085	9614571				
1404 C152015	3.8501475 8.2453	COURSE		mix 3 4		8471111 C152015	11405684				
1405 C16025	NA	000130				8502767 C16025	1091317:				
1406 C17766	NR 1					8520994 C17766	8967234				
1407 C19655	4.9487600	Diant pair 2 1	0/	Diant mix Onom	10	8590520 C19655	9455861				
1408 C17890	NA 1	PIQNI MIX 3 4	70		17	8686037 C17890	956100(
1409 C18155-COMBO	NA 1					MBO 8775785 C18155-COMBO	996367				
1410 C17622	101	Execution	00	Commorcial mix 3.8	4	8784887 C17622	938477;				
1411 C19751	5.3890719	EXCOVATION	72	Commercial mix 5 o	0	8817610 C19751	9838466				
1412 C17110	7.0892429					9132142 C17110	981919!				
1413 C17772	N/1 1	Unclassified				9178343 C17772	10047700				
1414 C13898	8.1143246		20		4	9232818 C13898	982222				
1415 C17988	104 1	Gen Asphalt	82	Plant mix 1 2	4	9278085 C17988	1034234				
1416 C10202	6.3189683		02			9371539 C16202	1000000				
1417 C13805	7.4151750 0.9250	commercial mix				9440193 C13808	1276453				
1418 (14986	5.8636312	COMMERCIALITIK				8952080 014900	12054080				
1419 (2030)	7.0527210					9550401 (20307	1084320.				
1420 014464	(1) 0.3353		17		4	9608134 C14404	10221498				
1421 0194764		Special borrow near	6/	Plant mix 3 8	4	9645000 0194704	10/05000				
1422 019128			· · · ·			9697696 019126	106/456:				
1423 (18070	5 2031100	line				9714475 C18070	10508411				
1424 017024	5.0021190					9/34/90 01/024	10488881				
1425 (176003	4 2626	Embankment in place	57	Commercial mix mm	2	9/55/82 01/6003	10260721				
1427 C18095-ALT	100 116 06 0	LINDUNKINGII III PIUCE	57	COMMERCIALITIK	5	T 0012862 C18095-ALT	1049100(
1420 C13070T-ALT A	7 7341				-	TA 0055160 C13070TALTA	11334604				
1420 (1815)	6.6227	Stad	41	Concrete class	2	10240773 C18151	1127234;				
1420 (17718	100 0.000	21661	41	CONCIENC CIUSS	2	10258301 017718	10000271				
1431 C18908-COMBO	NA			structuro		W80 10295964 C18908-COMBO	11669000				
1432 C13003	ALK 3.4011			SILUCIOLE		10335443 C13003	1187490;				
1433 C17735	N/A 8.472F	Execution borrow	20	Caparata alam daak	1	10341576 C17735	1152879!				
1434 C18367	NA 5.5254	EXCOVATION DOTTOW	JZ	Concrete class deck		10350253 C18367	11548550				
1435 C17169	7.5368972 6.9343					10544526 C17169	11337835				
1436 C17757	5.8036280	Concrata Class DD	21	Concrete class SD	1	10591493 C17757	12146261				
1437 C17136R-ALT	4.7706847		31	COUCIEIE CIUSS 3D		LT 10690902 C17136R-ALT	1222569:				
1438 C13166	NA	Dutation				10698459 C13166	10976195				
1439 C15790A	NA 7.8720	Bridde		repair		10785717 C15790A	1173442				
1440 C16222	NA 8.7688		0.1	. • • •		11001349 C16222	1375136!				
1441 C15504	NA 6.9087	Concrete General	31			11268135 C15504	1457820;				
1442 C20306	NA 7					11395807 C20306	12938338				
1443 C13579-ALT	6.0544395 4.2626	80 NA 5.883322 8.186130500000009 6.2982135 NA 6.542472	8.2200431999999992	5.6539969 6.33682580 7.642716 2.94685720 NA 5.9717717 14.54534 7.20	J3033 4.7369876 3 C13579-	ALT 10459946 C13579-ALT	14731462				
		AN ALEANAN ANALANA ANA			incontensione la terraria						
Showing 1.376 to 1.431 of 1.500 entries											
	and discourses										

Multivariate Linear Regression

MLR is similar to univariate linear regression, but instead of having one independent variable, there can be several estimators.

Our dependent or response variable is charge days and the independent variables are bid quantities, EE, locations, and project types.

> reg.gen<-lm(chgdays ~ sewer + perfpipe + pvc + concpipe + classd + pavmark + muckexc + rockexc +
concrete + concpav + strback10 + emban100 + asphalt10+ unclexc10 + strexc10+ asphre10 + agg10 + lnu
sd10, data = newrun3)</pre>

Unusual Observations.

##### Unusu	al Obervo	ations #	#########	########

#--> Outliers (Excessive Residuals, Y direction)

plot(rstudent(run19.lmo),pch=19)
outlierTest(run19.lmo)
influenceIndexPlot(run19.lmo)
marginalModelPlots(run19.lmo, id.n=4)

Plot the Studentized residuals in an Index plot

- # Actual statistical test for largest outlier
- # As before

#--> Leverage (Hat Values, X direction)

plot(hatvalues(run19.lmo)) # Plot of all Hat Values leveragePlots(run19.lmo, id.n = 4) # Specific Leverage plots by predictor (and fitted values)like avPlots() influenceIndexPlot(run19.lmo, id.n=4) # As before

#--> Influence (Y and X direction simultaneously)

(crit.cd <- qf(.5, 3,39, lower.tail = FALSE)) # Critical F for assessing Cook's Distances
plot(cooks.distance(medical2.lmo), type="b") # Plot of Cook's Distances
abline(h=crit.cd, col = "red", lwd=6) # Cut off line (critical values) for Cook's D
influenceIndexPlot(fuel2.lmo, id.n=4) # Again, as before
influencePlot(fuel2.lmo) # A new one and useful too</pre>

Variable Transformations



Variable Transformation (cont.)

More advanced method

Suggests the power transformation for a variable in presence of all other variables, which is better than older methods that transformed a variable without considering others.

Tranformations
#prodictor#

#predictor#

boxTidwell(chgdays ~ sewer + perfpipe + pvc + concpipe + classd + pavmark + muckexc + rockey

#response#

powerTransform(chgdays ~ sewer + perfpipe + pvc + concpipe + classd + pavmark + muckexc + re

MLR model for time estimation



*Based on a standardized Engineer's Estimate using the NHCCI index <u>https://www.fhwa.dot.gov/policyinformation/nhcci.cfm</u>

Validation



 $Y = 44.532 + 9.253E - 6 * X_1 + .008 * X_2 + .001X_3 + 5.421E - 5 * X_5 + .002 * X_5 + \varepsilon^1$ Goodness of fit F=121.354, significance =.000, Adjusted R²=.746, Mean Percent Error= 44.59%, Median Percent Error= 29.54%.

Predicted vs. Actual Durations (log scale)



Path Forward

 Taylor et al. (2013) developed a similar tool for Kentucky Transportation Cabinet.
 Creating a multi-state time determination tool (CO, GA, MS, MT)

Path Forward (cont.)

Develop a Machine Learning Estimation Model (analyzing Case-Based Reasoning and Neural Networks). Questions?