Football Player Weight Analysis Computer Lab

Canon City High School vs. Pueblo County High School

Mark Heinen

September 20, 2014

Table of Contents

I. Problem Statement

A. View Appendix A for problem specifics.

B. Summary of Problem Requirements

1. Create a single graph with both schools' histograms and cumulative distribution functions.

2. Derive each teams' Descriptive Statistics as provided in MS Excel's Data Analysis Add-In.

3. Provide a brief explanation of each parameter in the descriptive statistics.

4. Determine the following probabilities:

a) The weight of a CCHS player exceeds 175 pounds

b) The weight of a PCHS player is less than 150.

c) The weight of a PCHS player is more than 180 but less than 210 pounds

d) The probability a randomly selected PCHS player outweighs a randomly selected CCHS player.

e) Explain how to obtain the difference of 2 normal distributions and SHOW THIS in a sketch.

f) What conclusions can be made about the player weights and probability distributions from both schools?

II. Solution Technique

A. Sort both team's weights from lowest to highest weight.

B. Identify and analyze any outliers.

C. Derive descriptive statistics for each team independently.

D. Determine bin size for histogram development.

E. Use Excel's Data Analysis Tool (histogram function) to obtain histograms for each team.

F. Create a joint frequency table containing both team's data.

G. Create a bar chart showing both team's frequency and cumulative probabilities.

H. Use the descriptive statistics for each team and the standard normal (Z) curve/formulas to calculate required probabilities.

I. Research to obtain the difference of 2 normal distributions to calculate and create the curve(s) to represent this difference.

J. Draw conclusions about the player weights and probability distributions from both schools.

III. Data Analysis

A. Both team's weights are sorted from lowest to highest weight as shown in Figure 1. Two (2) outliers were identified in yellow. BOTH outliers will NOT be included in subsequent analysis.

Figure 2

Statistic	*CCHS*	*PCHS*
Mean	166.13	179.37
Standard Error	3.61	5.90
Median	171.93	179.55
Mode	#N/A	#N/A
Standard Deviation	27.77	45.31
Sample Variance	771.02	2052.75
Kurtosis	0.73	-0.02
Skewness	-0.46	0.36
Range	154	204
Minimum	86	92
Maximum	240	296
Sum	9802	10583
Count	59	59

B. Descriptive statistics for each team are shown independently in Figure 2.

.

C. The bin range and width for histogram development is selected as 100 to 300 pounds in 20 pound increments.

D. Excel's Data Analysis Tool (histogram function) was used to obtain histograms for each team as shown in Figure 3 and Figure 4.

Figure 3

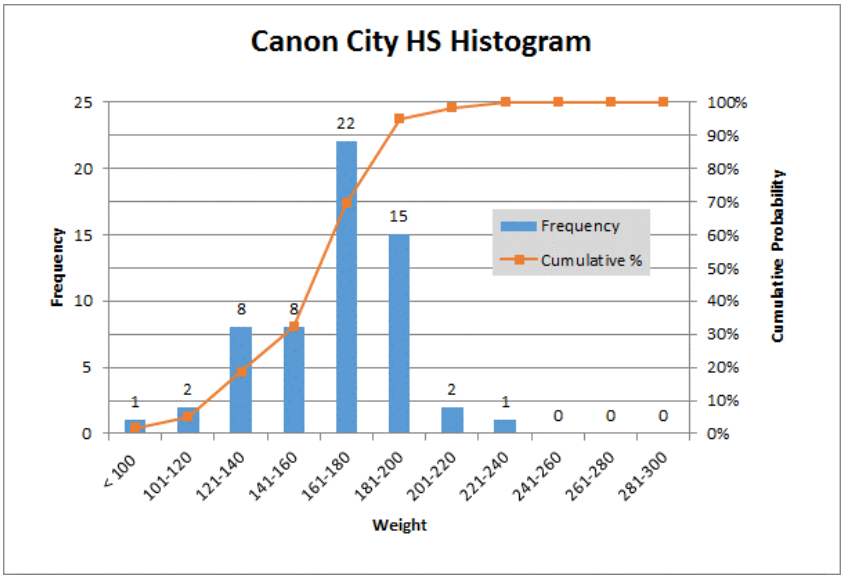
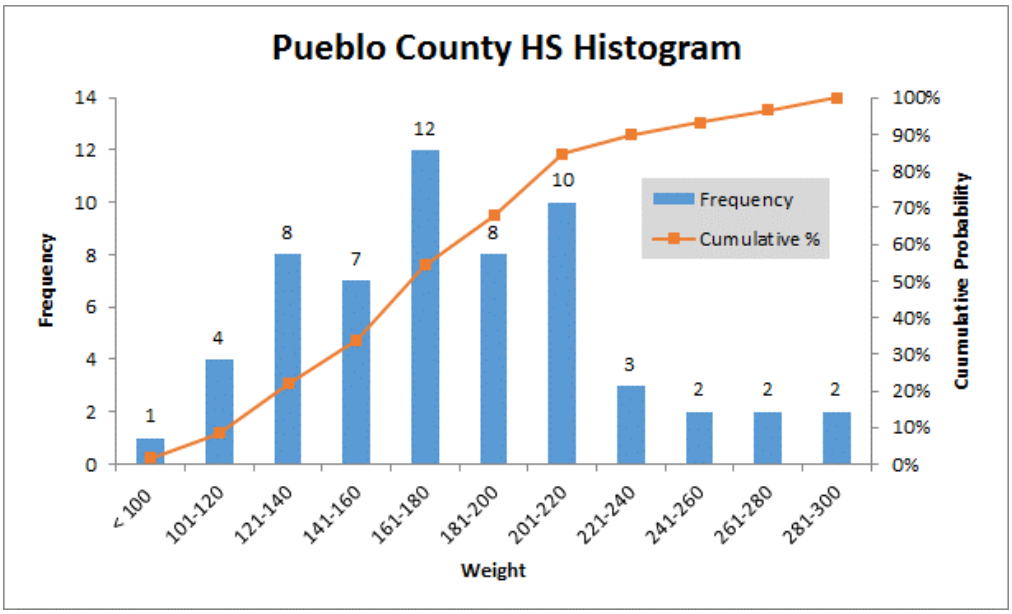


Figure 1

CCHS	PCHS
65	36
86	92
110	100
111	101
121	115
121	116
125	121
128	123
131	125
134	130
134	132
137	137
143	138
148	140
148	141
148	150
152	152
156	155
157	157
159	157
161	159
163	160
163	162
164	163
165	164
165	168
166	169
168	170
171	177
172	178
172	180
172	180
172	180
173	182
173	182
174	185
175	185
175	193
176	195
176	196
176	199
177	204
181	204
183	205
184	207
184	209
185	212
187	213
190	214
191	215
193	217
193	221
193	223
193	231
195	244
197	250
199	265
205	266
211	282
240	296

Figure 4



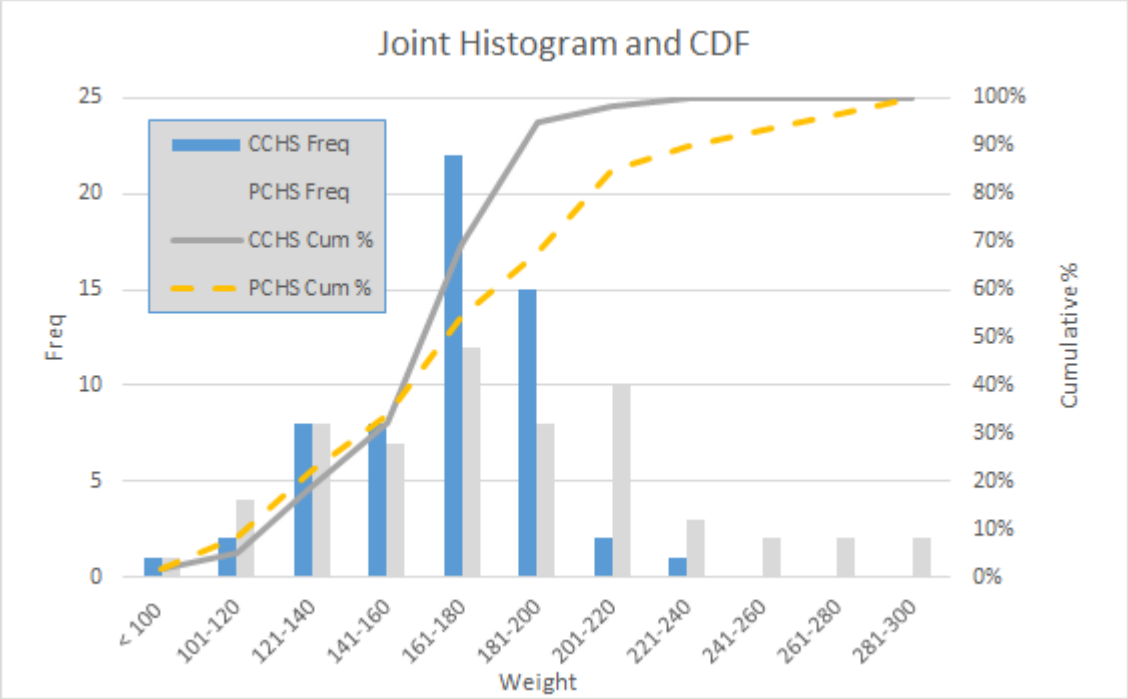
E. A joint frequency table for both teams is shown in Figure 5.

Figure 5

Bin	*CCHS Freq*	*PCHS Freq*	*CCHS Cum %*	*PCHS Cum %*
< 100	1	1	2%	2%
101-120	2	4	5%	8%
121-140	8	8	19%	22%
141-160	8	7	32%	34%
161-180	22	12	69%	54%
181-200	15	8	95%	68%
201-220	2	10	98%	85%
221-240	1	3	100%	90%
241-260	0	2	100%	93%
261-280	0	2	100%	97%
281-300	0	2	100%	100%

F. Figure 6 shows a bar chart showing both team's frequency and cumulative probabilities.

Figure 6



G. Use the descriptive statistics for each team and the standard normal (Z) curve / formulas to calculate required probabilities.

a) The probability density functions are shown in Figure 7 with the density functions below.

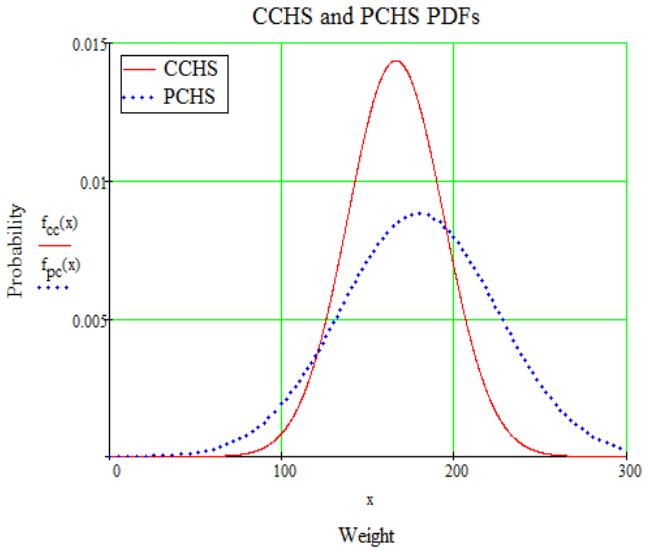
$\mu_{cc} := 166.13 \qquad \sigma_{cc} := 27.77$

$f_{cc}(x) := \text{dnorm}(x, \mu_{cc}, \sigma_{cc})$

$\mu_{pc} := 179.37 \qquad \sigma_{pc} := 45.31$

$f_{pc}(x) := \text{dnorm}(x, \mu_{pc}, \sigma_{pc})$

Figure 7



b) The probability the weight of a CCHS player exceeds 175 pounds is calculated below using both calculus and z values.

$$p_a := 1 - \int_0^{175} f_{cc}(x)\,dx = 0.375 \qquad z_a := \frac{175 - \mu_{cc}}{\sigma_{cc}} = 0.319$$

$$1 - \text{cnorm}(z_a) = 0.375$$

c) The weight of a PCHS player is less than 150 pounds is calculated below using both calculus and z values.

$$p_b := \int_0^{150} f_{pc}(x)\,dx = 0.258 \qquad z_b := \frac{150 - \mu_{pc}}{\sigma_{pc}} = -0.648$$

$$\text{cnorm}(z_b) = 0.258$$

d) The weight of a PCHS player is more than 180 but less than 210 pounds is calculated below using both calculus and z values.

$$p_c := \int_{180}^{210} f_{pc}(x)\,dx = 0.245$$

$$z_{cL} := \frac{180 - \mu_{pc}}{\sigma_{pc}} = 0.014 \qquad z_{cU} := \frac{210 - \mu_{pc}}{\sigma_{pc}} = 0.676$$

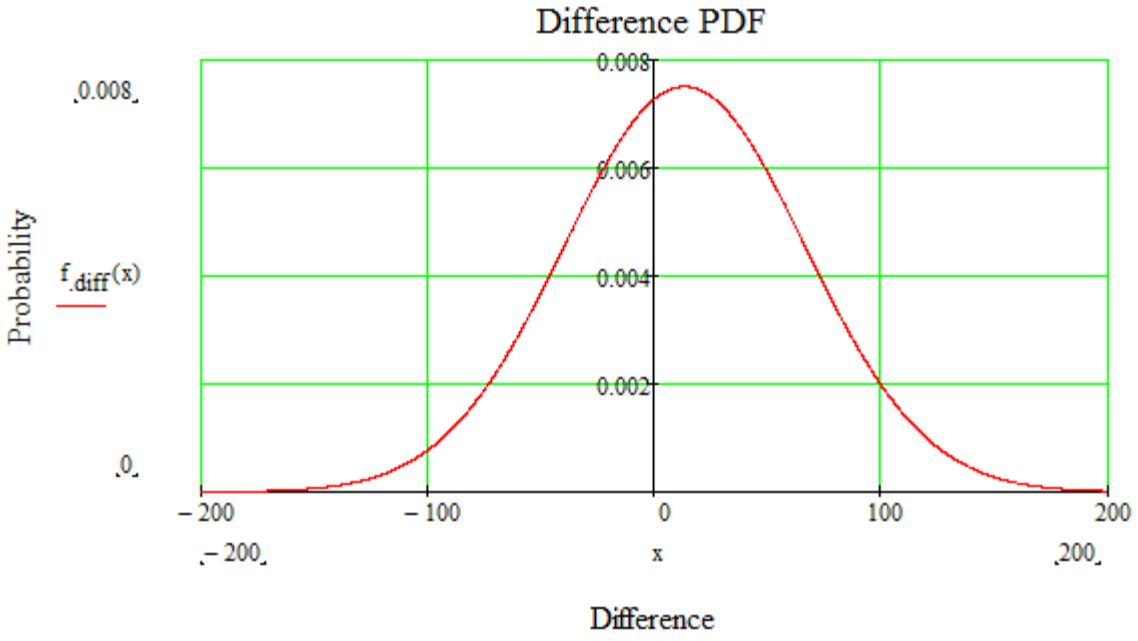
$$A_{cL} := \text{cnorm}(z_{cL}) = 0.506 \qquad A_{cU} := \text{cnorm}(z_{cU}) = 0.75 \qquad A_{cU} - A_{cL} = 0.245$$

H. Research to obtain the difference of 2 normal distributions to calculate and create the curve(s) to represent this difference.

$A_{cL} := cnorm(z_{cL}) = 0.506$ $A_{cU} := cnorm(z_{cU}) = 0.75$ $A_{cU} - A_{cL} = 0.245$

$\sigma_{diff} := \sqrt{\sigma_{cc}^2 + \sigma_{pc}^2} = 53.143$ $\mu_{diff} := \mu_{pc} - \mu_{cc} = 13.24$ $f_{diff}(x) := dnorm(x, \mu_{diff}, \sigma_{diff})$

Figure 8



The probability a PCHS player outweighs a CCHS player is the area to the right of x = 0. This area is computed using calculus and Z values as shown to the right. The area = 59.8 %.

$$\int_0^{200} f_{diff}(x)\,dx = 0.598$$

$$z_{diff} := \frac{0 - \mu_{diff}}{\sigma_{diff}} = -0.249$$

$A_{diff} := 1 - cnorm(z_{diff}) = 0.598$

I. Draw conclusions about the player weights and probability distributions from both schools.
   1. The mean weight of PCHS players exceeds the CCHS player's weights by 16 pounds.
   2. The PCHS player has greater variance.
   3. The probability that the weight of a randomly selected PCHS player exceeds the weight of a randomly selected CCHS Player is 59.8%.

IV. Conclusions

A. Probabilities:

1. The weight of a CCHS player exceeds 175 pounds = 37.5%

2. The weight of a PCHS player is less than 150 = 25.8%

3. The weight of a PCHS player is more than 180 but less than 210 pounds = 24.5%

4. The probability a randomly selected PCHS player outweighs a randomly selected CCHS player = 59.8%

B. Definitions

Statistic	Definition
Mean	
Standard Error	
Median	
Mode	
Standard Deviation	
Sample Variance	
Kurtosis	
Skewness	
Range	
Minimum	
Maximum	
Sum	
Count	

Appendix A

Use Excel to create separate histograms for each high school's player's weight. You select the bin (class) size and use Excel's histogram tool from their data analysis package. **You may download ALL player data as a file from: www.markeredwards.com. Some typical player data is shown to the left.**

Player (random selection)	CCHS	PCHS
1	205	239
2	230	295
3	200	192
4	134	265
5	165	209
6	175	162
7	145	121
8	161	109
9	107	140
10	187	164
11	160	226
12	113	175
13	120	119
14	176	200
15	160	138
16	150	169
17	173	167
18	121	164
19	152	247
20	133	224
21	186	143
22	185	99
23	132	209
24	105	205
25	151	180
26	171	145
27	185	149
28	152	166
29	143	197
30	171	185
31	121	155
32	166	203
33	178	218
34	111	108
35	110	159
36	142	199
37	118	161
38	177	226
39	135	251
40	181	154

download the data from the web site.

- Create a **single** graph with both schools histograms and cumulative distribution functions.
- Derive each teams' Descriptive Statistics as provided in MS Excel's Data Analysis Add-In.
- Provide a brief explanation of **each** parameter in the descriptive statistics (example to right.)

Example Only

Stat	CCHS	PCHS
Mean	166.131019	179.3707182
Standard Error	3.614985347	5.898507257
Median	171.926774	179.5452308
Mode	#N/A	#N/A
Standard Devia	27.76722933	45.30729878
Sample Varianc	771.0190244	2052.75087
Kurtosis	0.731259799	-0.01657672
Skewness	-0.457731309	0.357705248
Range	154.0583098	204.0750569
Minimum	85.83822738	92.3147687
Maximum	239.8965372	296.3898256
Sum	9801.730118	10582.87238
Count	59	59
SD(pop)	27.53090818	44.92169278
SD(sam)	27.76722933	45.30729878

- Determine the following probabilities:

1. the weight of a CCHS player exceeds 175 pounds
2. the weight of a PCHS player is less than 150.
3. the weight of a PCHS player is more than 180 but less than 210 pounds
4. the probability a randomly selected PCHS player outweighs a randomly selected CCHS player. How to do this??? Explain how to obtain the difference of 2 normal distributions and SHOW THIS in a sketch.
5. What conclusions can be made about the player weights and distributions from both schools?

Present your findings in a quality, MS Word document (or equivalent such as SUN Open Office). Include a cover page, Table of Contents, appendices, etc.

Submit your document as an e-mail attachment sent to mheinen_1@msn.com NLT midnight, 9-15-2014. Please confirm Mr. Heinen's receipt of the e-mail and its attachment.

Ensure the attached file is named as follows: CompLab-Players_LASTNAME_FIRSTINITIAL

Example: CompLab-Player_HEINEN_M.docx (the computer will add the suffix).