Analyzing the Seahawks Offensive Play-Calling During The 2013 Regular Season

Adam Brand

May 2014

Why Analyze Play Calling?

- Football most popular sport in U.S.
  - 205 mill U.S. viewers
  - 81% homes, 70% viewers
- Competitive Advantage

What Did I Analyze?

- All offensive plays in 2013 season
- Designed observational study
- Watched every play
- 1000+ play events & 50+ fields

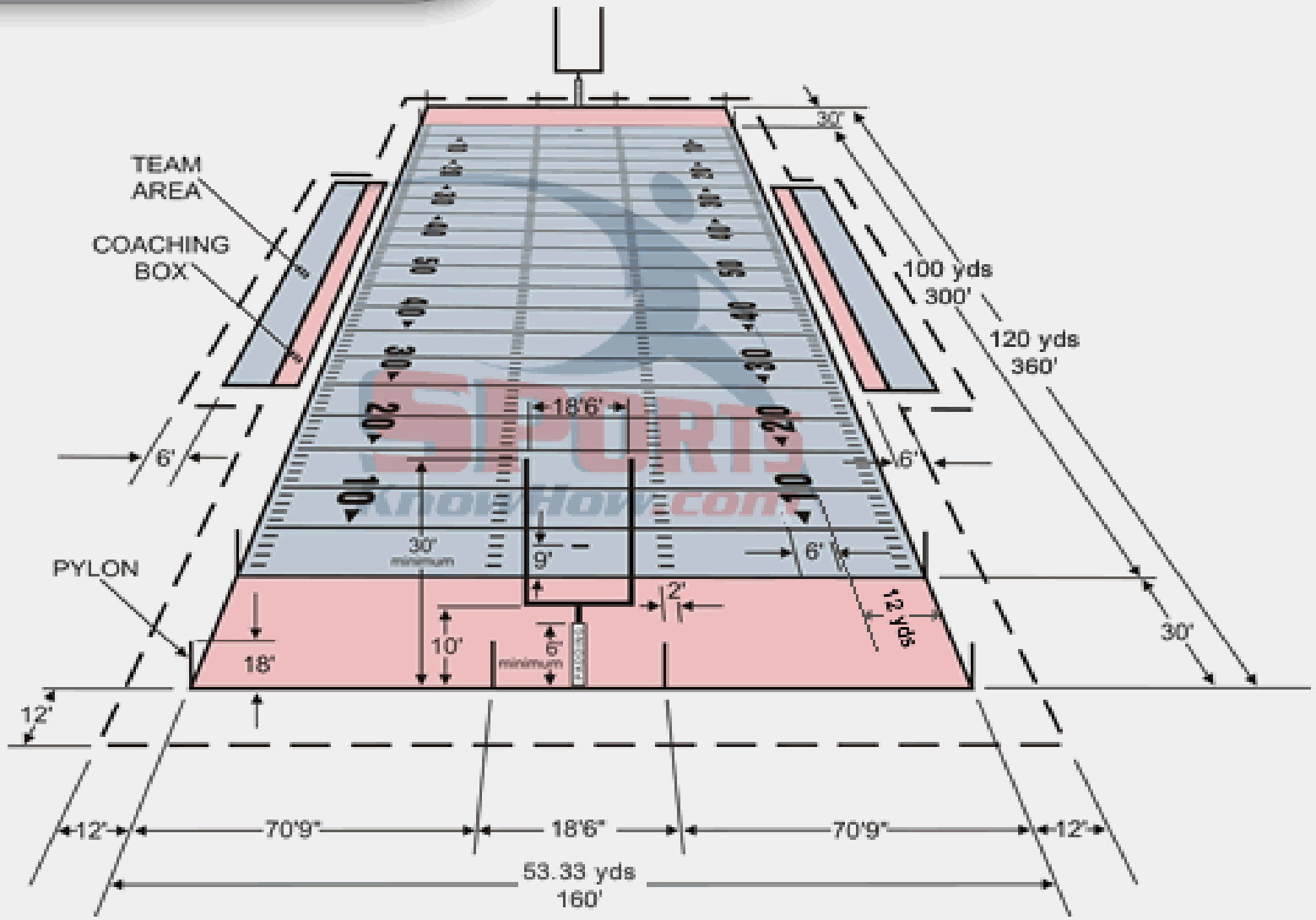
Which Questions Did I Ask?

- Play call preference?
- Play call and down dependent?
- Some plays work better?
- Probability of gaining Y yards?

SPORTS KnowHow.com

professional football field

diagram



Types of Plays

- Run:
  - Left, Mid, Right
- Pass:
  - Left, Mid, Right
- Scramble and Sack

Prefer Run vs. Pass?

Total Run Plays	451
Total pass plays	538
% running plays	0.456
% passing plays	0.544

Statistical significance?

- 2 Proportion Z Test

Conditions for Valid 2 Proportion Z Test

1. Samples randomly selected independently.

2. Both sample large

$n_1\hat{\rho}_1 \geq 15, n_1\hat{q}_1 \geq 15,$ and $n_2\hat{p}_2 \geq 15, n_2\hat{q}_2 \geq 15.$

Test Statistic

$$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sigma =_{(\hat{p}_1 - \hat{p}_2)}}$$

$$\boldsymbol{\sigma}_{(\hat{p}_1 - \hat{p}_2)} \approx \sqrt{\hat{p}\hat{q}\left(\frac{1}{n1} + \frac{1}{n2}\right)}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Result

- Significance level of 0.05

Ho: $p(pass) - p(run) \leq 0$

Ha: $p(pass) - p(run) > 0$

p-value = .00005

Answer: Yes, we prefer pass!

Interesting Result

- Sportscharts.com reports the Seahawks run the ball 54.79% of the time

- Ranked highest in the NFL

- Count scramble as run

Play Call Breakdown

Type of Play Called	Down 1	2	3	4	Total
Count of Run Right(# yds)	50	35	6		91
Count of Run Mid (# yds)	173	112	25	5	315
Count of Run Left (# yds)	16	24	5		45
Count of Pass Left	57	37	43	2	139
Count of Pass Mid	39	25	41		105
Count of Pass Right	71	67	53	1	192
Count of Scramble	24	16	16	2	58
Count of Sack (# yards lost)	15	11	17	1	44
Total	445	327	206	11	989

One-Way Chi-Squared Test on Play Call Type

Ho: p1 = p2 = p3 = p4 = p5 = p6 = 1/6

Ha: at least one proportion exceeds 1/6

Conditions Required for a Valid $\mathbf{X}^2$ Test: One-Way

1. A multinomial experiment has been conducted.

2. The sample size *n* will be large

Test Statistic

$$\chi^2 = \sum \frac{[n_i - E_i]^2}{E_i}$$

$$\chi^2 \text{ has } (r-1)df$$

Result – 6 Play Types

Hawks prefer some plays!

Using a sig level 0f .05 and degrees of freedom: 5

p – value ≈ 0

Run Direction Preference? Yes!

- 2 proportion Z – tests
- Result: Mid > Right > Left
- All p-values significant at .05 level of significance

Pass Direction Preference? Yes!

- 2 proportion Z – tests
- Result: Right > Left > Mid
- All p-values significant at .05 level of significance

Play Call and Down Dependent?

- Apply 2 Way Chi-Squared Test

Conditions for Valid X^2 Test: 2-Way

1. Multinomial experiment with *r* x *c* possible outcomes.

2. Large sample sizes

Observed Data

	Down			
Type of Play Called	1	2	3	Total
Count of Run Right(# yds)	50	35	6	91
Count of Run Mid (# yds)	173	112	25	310
Count of Run Left (# yds)	16	24	5	45
Count of Pass Left	57	37	43	137
Count of Pass Mid	39	25	41	105
Count of Pass Right	71	67	53	191
Total	406	300	173	879

Expected Values

	Down			
Type of Play Called	1	2	3	Total
Count of Run Right(# yds)	42.03	31.06	17.91	91
Count of Run Mid (# yds)	143.19	105.80	61.01	310
Count of Run Left (# yds)	20.78	15.36	8.86	45
Count of Pass Left	63.28	46.76	26.96	137
Count of Pass Mid	48.50	35.84	20.67	105
Count of Pass Right	88.22	65.19	37.59	191
Total	406	300	173	879

2-Way Chi-Squared Test Statistic

$$\chi 2 = \sum \frac{\left[n_{\mathrm{ij}} - \hat{\mathrm{E}}_{\mathrm{ij}}\right]^2}{\hat{\mathrm{E}}_{\mathrm{ij}}}$$

$$\hat{\mathrm{E}}_{\mathrm{ij}} = \frac{R_{\mathrm{i}} C_{\mathrm{j}}}{n}$$

$$\chi 2_{\mathrm{x\ has}} (r-1)(c-1) df$$

Play Call and Down Dependent? Yes!

- 2-Way Chi-Squared Test with 10 degrees of Freedom at sig level .05

p – value ≈ 0

Observed Probability of Plays by Down

	Down		
Play type	1st	2nd	3rd
Count of Run Right(# yds)	0.11	0.11	0.03
Count of Run Mid (# yds)	0.39	0.34	0.12
Count of Run Left (# yds)	0.04	0.07	0.02
Count of Pass Left	0.13	0.11	0.21
Count of Pass Mid	0.09	0.08	0.20
Count of Pass Right	0.16	0.20	0.26
Count of Scramble	0.05	0.05	0.08
Count of Sack (# yards lost)	0.03	0.03	0.08

How Do Different Plays Perform?

- Analyze net yards gained (NYG)
- Compare means
- Fit probability distributions
- Estimated NYG

Mean Net Yards Gained by Play Type

Play Type Means	
Run Left	5.13
Run Mid	3.37
Run Right	5.84
Pass Left	9.38
Pass Mid	6.57
Pass Right	8.41
Sack	-5.91
Scramble	7.31

Run NYG vs. Pass NYG

	Run	Pass
Sample size	451	538
Mean	4.05	6.99
Std Dev	5.96	11.75

- 2 Sample Z Test

Conditions for Valid 2 Sample Z Test

1. Random Independent samples

2. Large sample sizes (> 30)

2 Sample Mean Test Statistic:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sigma\,(\bar{x}_1 - \bar{x}_2)}$$

$$\sigma\,(\bar{x}_1 - \bar{x}_2) \approx \sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}$$

Pass NYG > Run NYG? Yes!

Ho: mean pass – mean run ≤ 0
Ha: mean pass – mean run > 0

- Sig level of .05
- p – value ≈ 0

How About the Means of Specific Play Types?

Play Type Performance

Play Type	Sample	Mean	Std Dev
Run Left	45	5.13	5.83
Run Mid	315	3.37	4.82
Run Right	91	5.84	8.63
Pass Left	139	9.38	14.45
Pass Mid	105	6.57	7.93
Pass Right	192	8.41	11.69
Sack	44	-5.91	3.05
Scramble	58	7.31	8.29

Statistical Ranking

Play Type	Rank	P-value
Pass Left	1	0.257
Pass Right	1	0.212
Scramble	1	0.29
Pass Mid	1/2	0.27
Run Right	1/2	0.286
Run Left	1/2	0.027
Run Mid	7	0+
Sack	8	

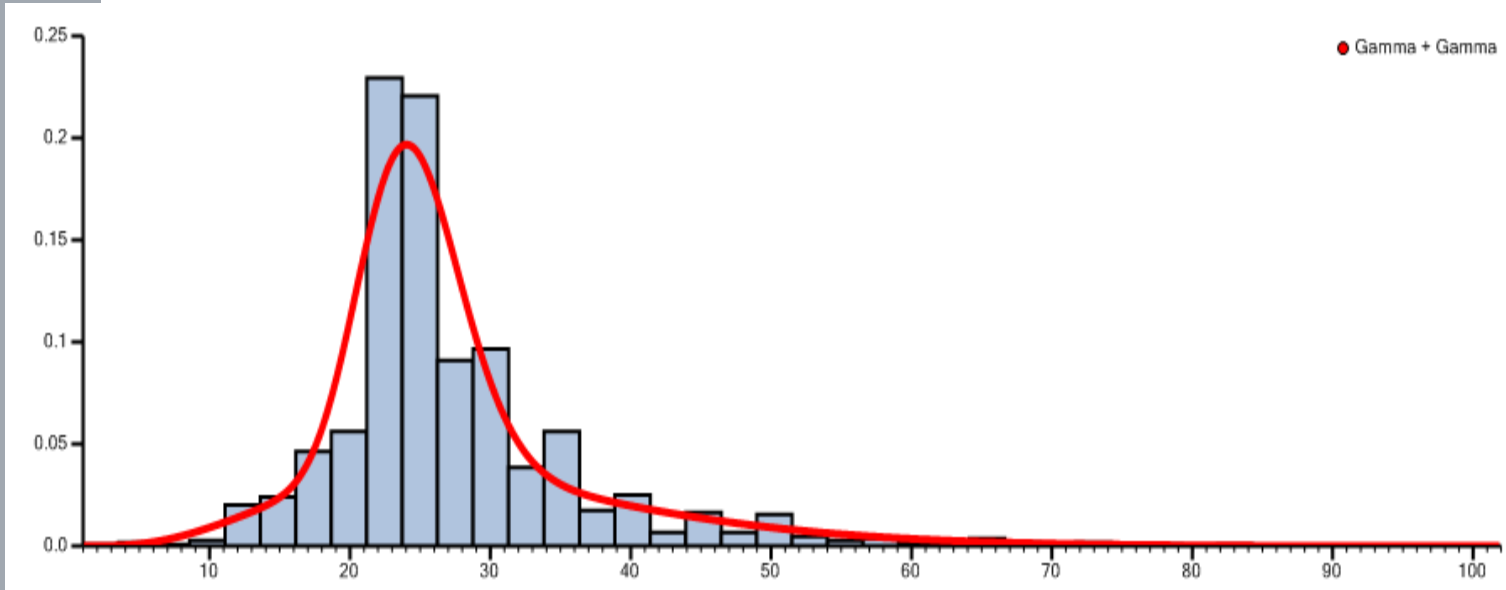
Modeling Play Performance

- Model NYG using AMOOF 3
- Y = r.v. for NYG
- Z = Y + 22
- Models fit Z

Aggregate NYG

Net Yards Gained	
Mean:	5.2234
Median:	3
Mode:	0
Std Dev:	9.532

Aggregate NYG



Z approximately follows: f(z) = .4231*Gamma(alpha = 5.253, theta = 5.8644) + .5769*Gamma(alpha = 46.447, theta = .53)

PDF for Aggregate NYG Model

$$f(z) = \frac{.4253 \cdot (z)^{4.253} \cdot e^{\left(\frac{-z}{5.8644}\right)}}{\Gamma(5.253)(5.8644)^{5.253}}$$

$$+ \frac{.5769(z)_{45.447} \cdot e - \left(\frac{z}{.53}\right)}{\Gamma(46.447)(.53)^{46.447}}$$

Aggregate NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

	1%	2%	3%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
VAR PDF	10.2073	12.2931	13.7913	15.9433	18.6655	20.0621	21.0571	21.8738	22.5968	23.2688	23.9162	24.5586	25.2135
VAR Data	12.0000	12.0000	14.0000	16.0000	19.0000	21.0000	22.0000	22.0000	22.0000	23.0000	24.0000	24.0000	25.0000
% Diff	14.9393	-2.4422	1.4911	0.3541	1.7607	4.4661	4.2860	0.5738	-2.7128	-1.1688	0.3492	-2.3273	-0.8542
L CTE PDF	8.1947	9.7601	10.8646	12.4966	15.0018	16.4730	17.4998	18.2950	18.9526	19.5217	20.0307	20.4980	20.9367
L CTE Data	8.4000	10.2000	11.1290	12.7647	15.0097	16.7484	18.0291	18.8295	19.3613	19.7562	20.2058	20.6301	21.0445
% Diff	2.4437	4.3128	2.3758	2.1003	0.0529	1.6445	2.9361	2.8386	2.1107	1.1874	0.8668	0.6402	0.5124

Aggregate NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%	99%
VAR PDF	25.2135	25.9001	26.6420	27.4743	28.4555	29.6986	31.4594	34.3511	39.2104	47.0361	52.2935	56.2553	62.7106
VAR Data	25.0000	26.0000	26.0000	27.0000	29.0000	30.0000	32.0000	35.0000	38.0000	46.0000	50.0000	55.0000	65.0000
% Diff	-0.8542	0.3844	-2.4692	-1.7566	1.8777	1.0047	1.6895	1.8539	-3.1853	-2.2524	-4.5869	-2.2824	3.5221
R CTE PDF	20.9367	21.3564	21.7654	22.1718	22.5845	23.0154	23.4843	24.0306	24.7287	25.6775	26.1680	26.4536	26.7838
R CTE Data	21.0445	21.4401	21.8226	22.2173	22.6252	23.0748	23.5623	24.1241	24.7860	25.6762	26.1257	26.4077	26.7253
% Diff	0.5124	0.3908	0.2619	0.2045	0.1799	0.2577	0.3310	0.3877	0.2311	-0.0050	-0.1618	-0.1738	-0.2188

Aggregate NYG

Mean and Standard Deviation

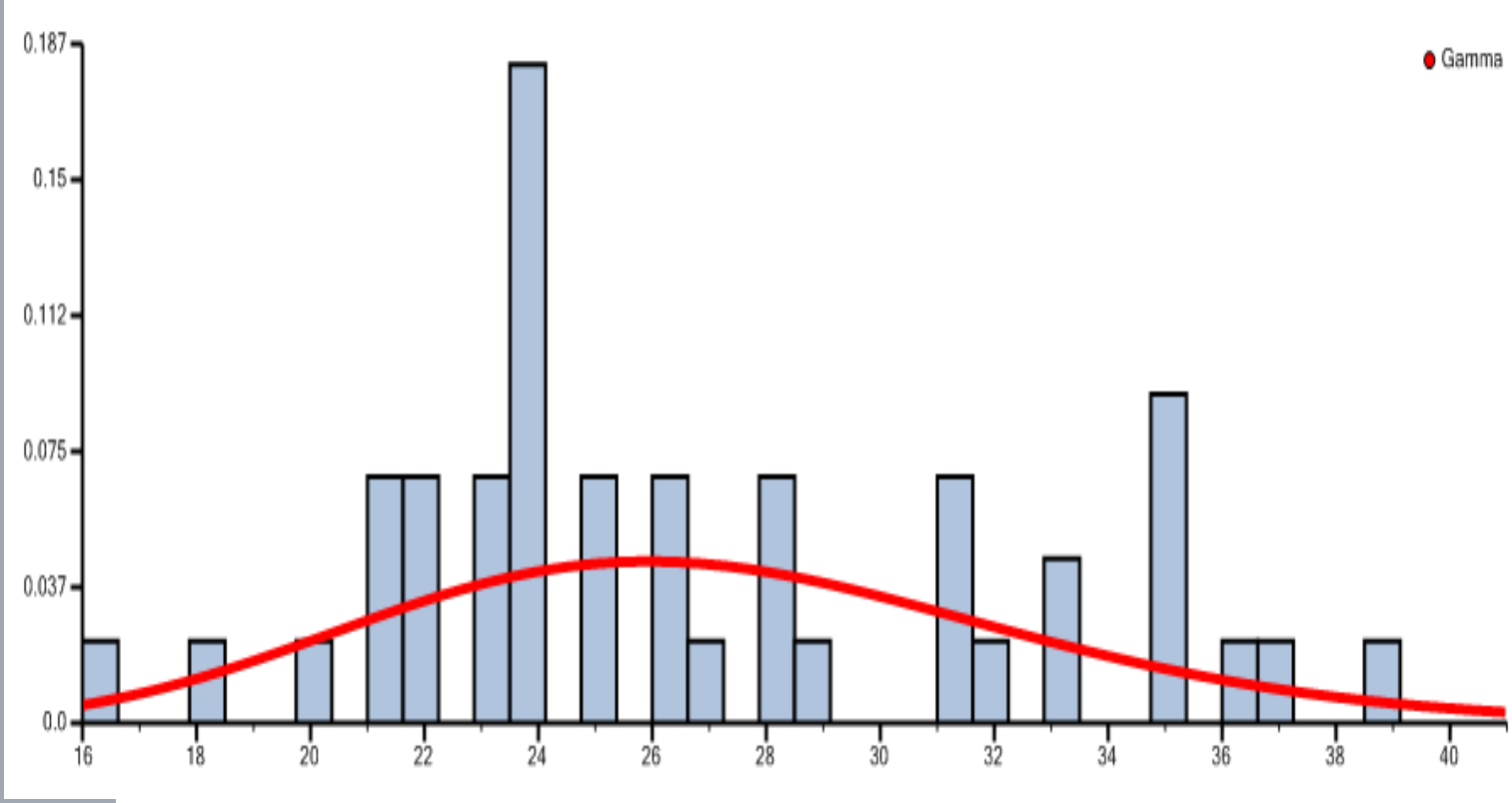
Net Yards Gained	
Mean	5.2234
Std Dev:	9.532

Implied Mean and Standard Deviation

Z Distribution	
Mean:	27.229
Std Dev:	9.659

Implied Net Yards Gained	
Mean:	5.229
Std Dev:	9.659

Run Left NYG



Gamma (alpha = 22.808, theta = 1.19)

Run Left NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

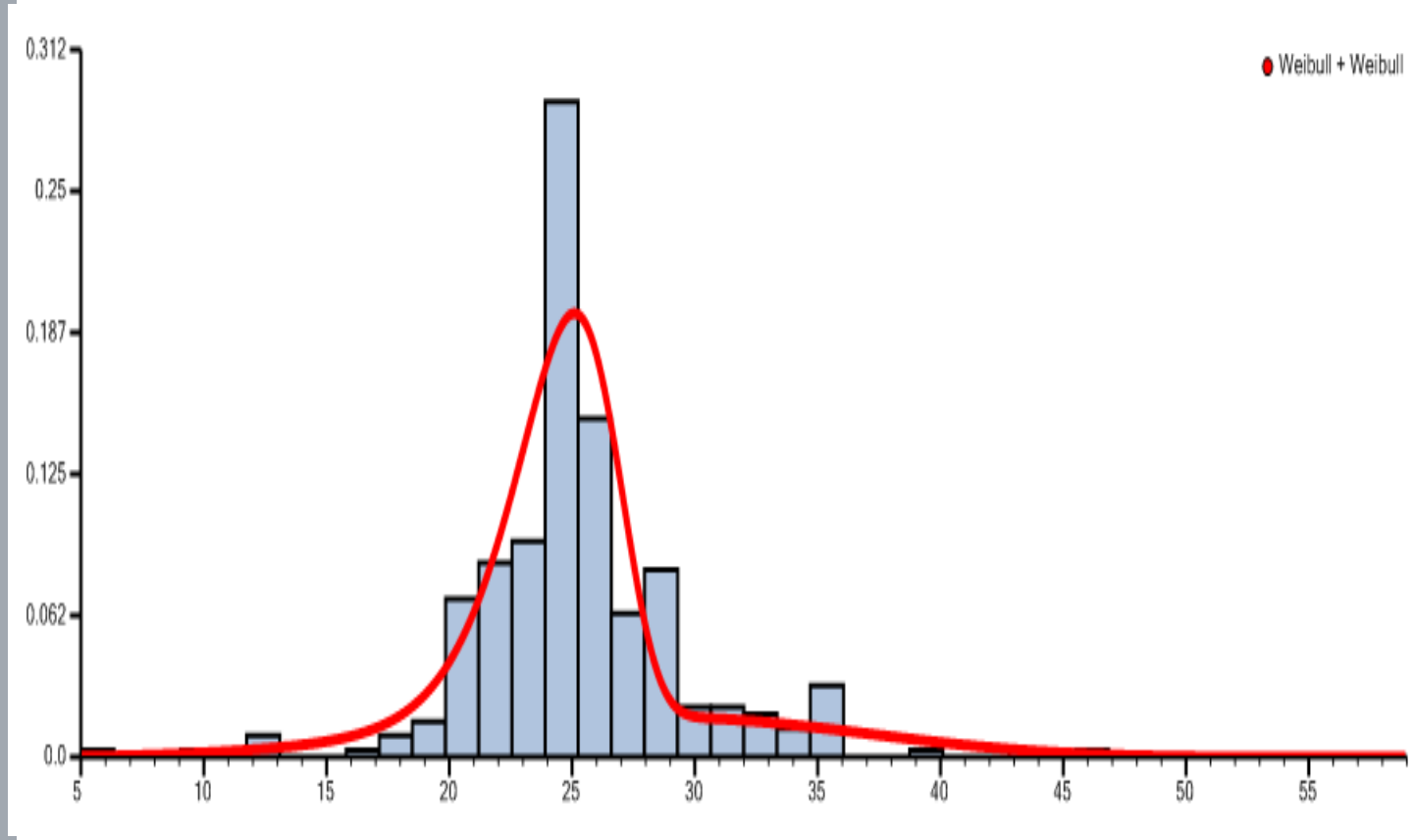
	1%	2%	3%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
VAR PDF	15.6835	16.7758	17.4956	18.5116	20.1543	21.3172	22.2734	23.1165	23.8919	24.6258	25.3361	26.0362	26.7378
VAR Data	16.0000	16.0000	18.0000	20.0000	21.0000	22.0000	23.0000	23.0000	24.0000	24.0000	24.0000	25.0000	25.0000
% Diff	1.9784	-4.8489	2.8020	7.4419	4.0272	3.1037	3.1592	-0.5066	0.4506	-2.6074	-5.5670	-4.1450	-6.9513
L CTE PDF	14.4277	15.3535	15.9538	16.7864	18.0918	18.9810	19.6876	20.2905	20.8268	21.3175	21.7756	22.2101	22.6278
L CTE Data	NaN	NaN	16.0000	17.0000	18.7500	19.5000	20.3333	20.8182	21.2308	21.6000	22.0000	22.2000	22.4545
% Diff	NaN	NaN	0.2890	1.2564	3.5106	2.6616	3.1757	2.5347	1.9025	1.3079	1.0200	-0.0457	-0.7715

Run Left NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%	99%
VAR PDF	26.7378	27.4519	28.1904	28.9675	29.8016	30.7195	31.7632	33.0090	34.6214	37.1044	38.7773	40.0342	42.0679
VAR Data	25.0000	26.0000	28.0000	28.0000	31.0000	31.0000	33.0000	35.0000	35.0000	37.0000	39.0000	41.0000	41.0000
% Diff	-6.9513	-5.5844	-0.6802	-3.4554	3.8657	0.9050	3.7478	5.6887	1.0818	-0.2822	0.5709	2.3557	-2.6046
R CTE PDF	22.6278	23.0337	23.4325	23.8281	24.2245	24.6264	25.0390	25.4699	25.9312	26.4477	26.6834	26.8128	26.9556
R CTE Data	22.4545	22.7083	23.1111	23.4483	23.7742	24.2121	24.8611	25.3421	25.8250	26.2857	26.5349	26.8182	26.8182
% Diff	-0.7715	-1.4330	-1.3905	-1.6197	-1.8943	-1.7110	-0.7156	-0.5044	-0.4112	-0.6162	-0.5597	0.0199	-0.5124

Run Mid NYG



f(z) = .718*Weibull(tau = 12.69, theta = 25.27)
+ .282* Weibull(tau = 3.77, theta = 30.83)

Run Mid NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

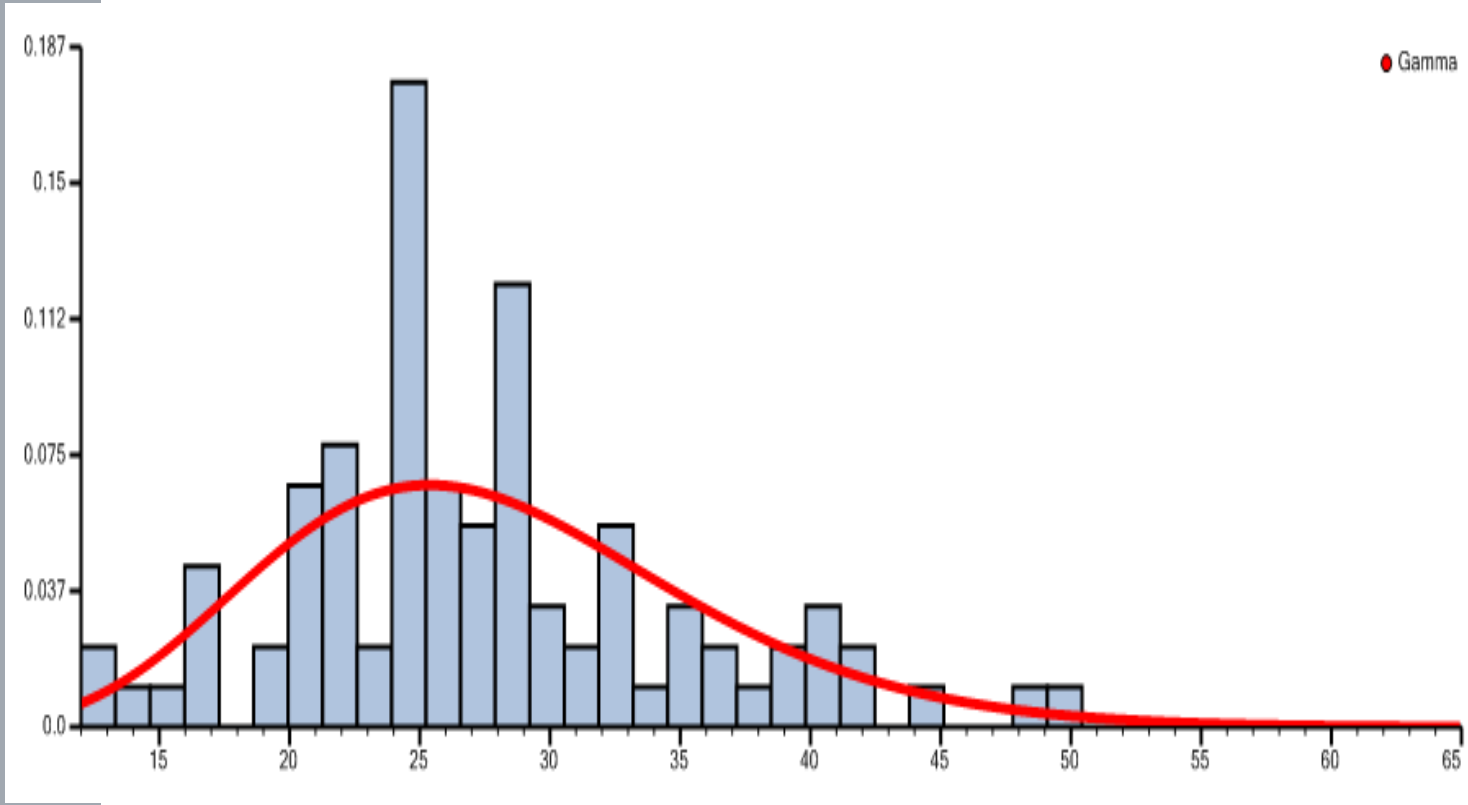
	1%	2%	3%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
VAR PDF	12.7312	15.1842	16.6569	18.3839	20.4141	21.4911	22.2394	22.8254	23.3170	23.7488	24.1413	24.5082	24.8596
VAR Data	12.0000	18.0000	19.0000	20.0000	21.0000	22.0000	23.0000	23.0000	24.0000	24.0000	24.0000	24.0000	25.0000
% Diff	-6.0934	15.6435	12.3323	8.0804	2.7902	2.3132	3.3072	0.7593	2.8459	1.0466	-0.5889	-2.1173	0.5617
L CTE PDF	10.0626	12.0710	13.3719	15.0672	17.3043	18.5338	19.3711	20.0054	20.5174	20.9487	21.3236	21.6572	21.9600
L CTE Data	9.0000	11.3333	13.5556	15.8000	18.0968	19.3191	20.0000	20.5769	21.0000	21.4364	21.7619	22.0000	22.2994
% Diff	-11.8065	-6.5092	1.3547	4.6382	4.3790	4.0652	3.1443	2.7776	2.2983	2.2748	2.0141	1.5580	1.5219

Run Mid NYG Analysis

Data Set Conditional Tail Expectations and Value at Risks

	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%	99%
VAR PDF	24.8596	25.2044	25.5513	25.9110	26.2986	26.7412	27.3026	28.2324	31.1399	35.6629	38.1939	39.9152	42.4542
VAR Data	25.0000	25.0000	26.0000	26.0000	26.0000	27.0000	28.0000	29.0000	30.0000	34.0000	36.0000	36.0000	40.0000
% Diff	0.5617	-0.8174	1.7257	0.3424	-1.1483	0.9585	2.4906	2.6470	-3.7997	-4.8909	-6.0942	-10.8757	-6.1356
R CTE PDF	21.9600	22.2393	22.5008	22.7492	22.9886	23.2236	23.4601	23.7101	24.0281	24.5152	24.7698	24.9152	25.0784
R CTE Data	22.2994	22.5491	22.7831	23.0196	23.2364	23.4449	23.6746	23.9401	24.2473	24.6689	24.8689	24.9773	25.0836
% Diff	1.5219	1.3741	1.2390	1.1749	1.0662	0.9439	0.9061	0.9607	0.9041	0.6228	0.3983	0.2487	0.0208

Run Right NYG



Gamma Distribution (alpha = 11.48, theta = 2.42)

Run Right NYG

Data Set Conditional Tail Expectations and Value at Risks

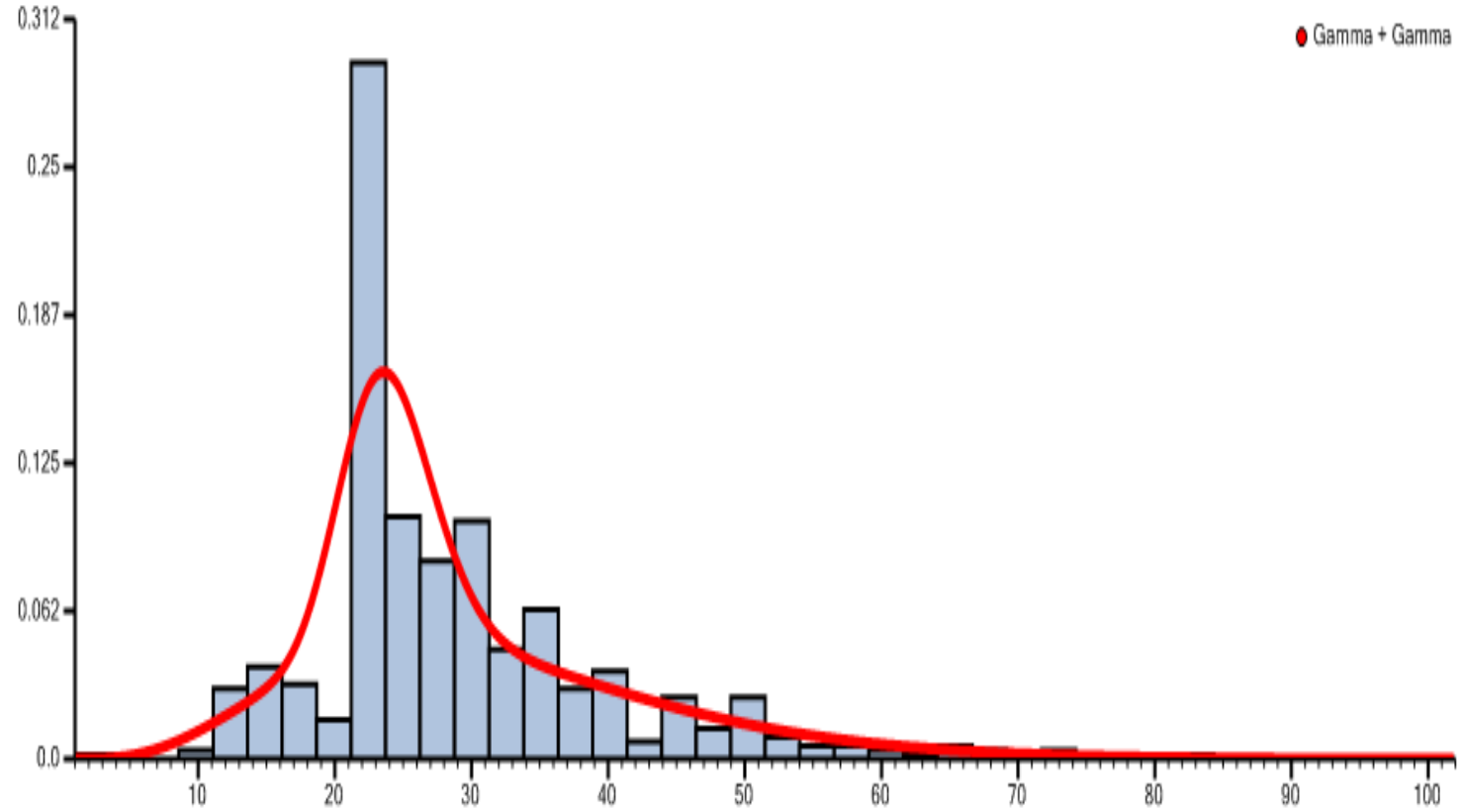
	1%	2%	3%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
VAR PDF	12.3292	13.6571	14.5502	15.8338	17.9618	19.5045	20.7937	21.9451	23.0154	24.0383	25.0369	26.0293	27.0313
VAR Data	12.0000	12.0000	14.0000	16.0000	19.0000	20.0000	22.0000	22.0000	24.0000	24.0000	25.0000	25.0000	26.0000
% Diff	-2.7432	-13.8089	-3.9302	1.0386	5.4643	2.4777	5.4834	0.2496	4.1025	-0.1596	-0.1476	-4.1170	-3.9666
L CTE PDF	10.8677	11.9592	12.6813	13.7008	15.3382	16.4796	17.4007	18.1961	18.9110	19.5708	20.1918	20.7853	21.3596
L CTE Data	NaN	12.0000	12.0000	13.2500	15.4444	16.7692	18.0000	18.7273	19.5556	20.1290	20.6944	21.1250	21.6444
% Diff	NaN	0.3403	-5.6775	-3.4020	0.6876	1.7273	3.3296	2.8362	3.2960	2.7731	2.4288	1.6081	1.3159

Run Right NYG

Data Set Conditional Tail Expectations and Value at Risks

	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%	99%
VAR PDF	27.0313	28.0588	29.1292	30.2637	31.4905	32.8508	34.4103	36.2884	38.7448	42.5798	45.1964	47.1785	50.4134
VAR Data	26.0000	27.0000	28.0000	29.0000	30.0000	32.0000	33.0000	36.0000	39.0000	42.0000	49.0000	50.0000	65.0000
% Diff	-3.9666	-3.9215	-4.0329	-4.3575	-4.9684	-2.6587	-4.2737	-0.8012	0.6543	-1.3804	7.7624	5.6431	22.4409
R CTE PDF	21.3596	21.9217	22.4774	23.0322	23.5919	24.1629	24.7534	25.3747	26.0455	26.8049	27.1551	27.3487	27.5636
R CTE Data	21.6444	22.1400	22.5370	23.0508	23.4286	23.9412	24.4167	25.0649	25.6667	26.5349	26.9205	27.1685	27.4222
% Diff	1.3159	0.9860	0.2648	0.0810	-0.6970	-0.9261	-1.3789	-1.2357	-1.4760	-1.0177	-0.8717	-0.6632	-0.5156

All Pass Plays NYG



Gamma-Gamma Model: f(z) = .386*Gamma(alpha = 50.84, theta = .472) + .614*Gamma(alpha = 5.205, theta = 6.144)

ALL PASS PLAYS NYG

Data Set Conditional Tail Expectations and Value at Risks:

	1%	2%	3%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
VAR PDF	9.5637	11.4573	12.8077	14.8338	17.9231	19.5972	20.7574	21.6985	22.5334	23.3188	24.0919	24.8834	25.7261
VAR Data	12.0000	12.0000	12.0000	15.0000	18.0000	22.0000	22.0000	22.0000	22.0000	22.0000	22.0000	25.0000	26.0000
% Diff	20.3027	4.5224	-6.7305	1.1079	0.4275	10.9218	5.6482	1.3704	-2.4243	-5.9947	-9.5087	0.4662	1.0533
L CTE PDF	7.7044	9.1427	10.1487	11.6393	14.0995	15.6751	16.8071	17.6937	18.4318	19.0741	19.6530	20.1899	20.7008
L CTE Data	7.4000	9.7000	10.5625	11.7692	14.0189	16.2500	17.7009	18.5672	19.1429	19.5532	19.8605	20.2727	20.7881
% Diff	-4.1134	5.7451	3.9177	1.1038	-0.5755	3.5381	5.0497	4.7043	3.7147	2.4500	1.0445	0.4086	0.4200

ALL PASS PLAYS NYG

Data Set Conditional Tail Expectations and Value at Risks:

	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%	99%
VAR PDF	25.7261	26.6628	27.7585	29.1224	30.9318	33.3781	36.4922	40.3036	45.2514	52.9424	58.2011	62.1997	68.7618
VAR Data	26.0000	27.0000	28.0000	30.0000	31.0000	33.0000	36.0000	39.0000	45.0000	51.0000	57.0000	65.0000	70.0000
% Diff	1.0533	1.2490	0.8626	2.9252	0.2201	-1.1456	-1.3672	-3.3425	-0.5587	-3.8086	-2.1072	4.3081	1.7689
R CTE PDF	20.7008	21.1993	21.6988	22.2152	22.7699	23.3917	24.1099	24.9464	25.9297	27.1292	27.7116	28.0418	28.4164
R CTE Data	20.7881	21.2983	21.8447	22.4126	22.9814	23.6055	24.2860	25.0481	25.9628	27.1096	27.6123	27.9696	28.3289
% Diff	0.4200	0.4648	0.6678	0.8807	0.9202	0.9054	0.7254	0.4063	0.1275	-0.0724	-0.3595	-0.2580	-0.3089

Mixed Distributions

$$P(Z = Z)$$
$$= W_1 \cdot fx.(z) + W_2\, fx_2\,(z)$$

$$P(Z > Z)$$
$$= \int_Z^{00} (W_1\, fx_1(u) + w_2\, fx_2(u))du$$

$$= w_1 \int_z^{00} fx_1(u)du + w_2 \int_z^{00} fx_2(u)du$$

$$= w_1.(1 - Fx_1(z)) + w_2.(1 - Fx_2(z))$$

Probabilities of gaining AT LEAST Y Yards!

Z	Y	All	Run Left	Run Mid	Run Right	Pass
23	1	0.6709	0.7576	0.7328	0.6986	0.6694
24	2	0.5941	0.6933	0.6181	0.6496	0.6048
25	3	0.5168	0.6243	0.4792	0.5994	0.5417
26	4	0.4436	0.5532	0.3376	0.5489	0.4837
27	5	0.3782	0.4821	0.2245	0.4990	0.4327
28	6	0.3224	0.4133	0.1588	0.4502	0.3895
29	7	0.2767	0.3485	0.1298	0.4033	0.3534
30	8	0.2402	0.2892	0.1145	0.3588	0.3234
31	9	0.2113	0.2363	0.1016	0.3170	0.2981
32	10	0.1884	0.1901	0.0892	0.2783	0.2762
33	11	0.1701	0.1507	0.0774	0.2427	0.2567
34	12	0.1548	0.1177	0.0664	0.2104	0.2390
35	13	0.1418	0.0907	0.0562	0.1813	0.2226
36	14	0.1304	0.0689	0.0469	0.1553	0.2072
37	15	0.1200	0.0517	0.0386	0.1323	0.1926

Sources

SportingIntelligence.com

TVBytheNumbers.com (via Network TV Press Release)

SportsCharts.com

Special Thanks:

AMOOF 3 Team Here at CWU!