Analyzing NFL receiving touchdowns using negative binomial model

Summary

The propensity of NFL receiving touchdown scoring is highly heterogenous, a finding that has held up historically.

Background

Receiving in NFL football is mainly defined by three statistics: receptions, yards, and touchdowns. Traditionally, of the three, touchdowns are the most unreliable and seem to be the most "random" from season to season; we examine individual NFL players' underlying propensities to score touchdowns and investigate historical statistics in the context of our model.

The story

To model this process, we imagine each individual NFL pass-catcher spinning their own "lambda wheel" (Poisson distribution) at the beginning of the season to determine the number of touchdown receptions they will have. We know that there are many factors influencing this outcome (notably, targets and snaps played), but in our story all these variables are encompassed in an individual's "propensity" to score touchdowns – if a receiver is going to see the field and get thrown to more, it is something about the individual that is causing teams/coaches to give him increased chances. Treating touchdowns as a proportion of players' opportunities to score also resembles more of a choice-type process, which we are not interested in.

Models

We decide to fit a basic NBD model on our data on NFL receiving touchdowns using MLE as we are given all the raw counts; we will also build our NBD model using summary statistics to test the robustness, in addition to building 0-missing models for further insight on underlying propensities.

More than the skill itself, NFL receivers each have wildly different levels of opportunity, and thus I was expecting there to be high population heterogeneity; at the same time, I was somewhat hoping that the model would indicate low heterogeneity and that subsequently high touchdown numbers were largely flukes.

<u>Defining the data</u>

We use data from the 2021 NFL regular season for our models. Immediately, the question arises of what subset of players to use: are we to include defensive players, who will likely never catch a touchdown pass? Should we include practice squad players that did not play a single snap in the regular season?

We tentatively define our population as **NFL wide receivers, tight ends, and running backs that played in at least one game in the 2021 regular season** and will run our model on different subgroups to confirm this.

When we think of pass-catchers in football, wide receivers and tight ends are traditionally the positions that come to mind; however, as in modern football we have seen running backs increasingly utilized in the pass game, we include them in our analysis. Although offensive linemen and quarterbacks do sporadically catch passes, they do not fit into our story as their role in the game is altogether different – this could be addressed by implementing a spike, which would “define” who our zeros are for us.

One possible consideration was to only select players with at least one recorded target, but this excluded players that ran routes but were never thrown to, individuals we still wanted to keep in our model. Only considering players that caught a touchdown was another consideration, which we analyze later using a truncated NBD model.

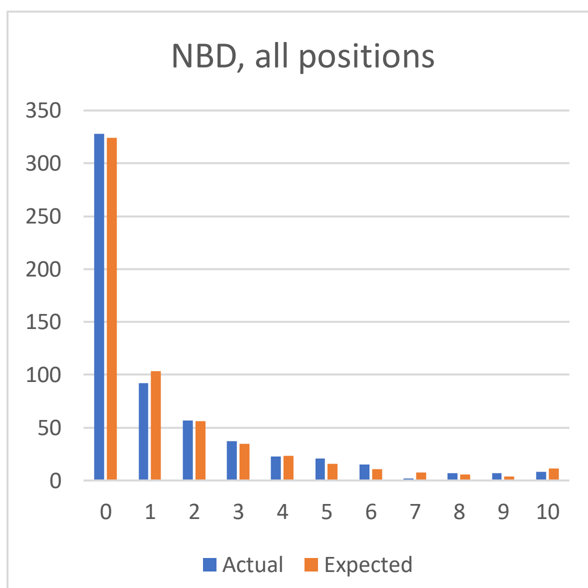
Our 2021 receiving datasets were pulled from The Football Database, a site that did not leave out players without targets or touchdown receptions (as opposed to NFL or ESPN); I scraped separate datasets for all positions, just WR/TE/RB, and then just WR.

For historical comparison, I also pulled NFL receiving data from Pro Football Reference from 2007 and 1987 – the current single season touchdown records are held by Randy Moss (23) in 2007, and Jerry Rice (22) in 1987 – and manually filtered out quarterbacks and offensive linemen.

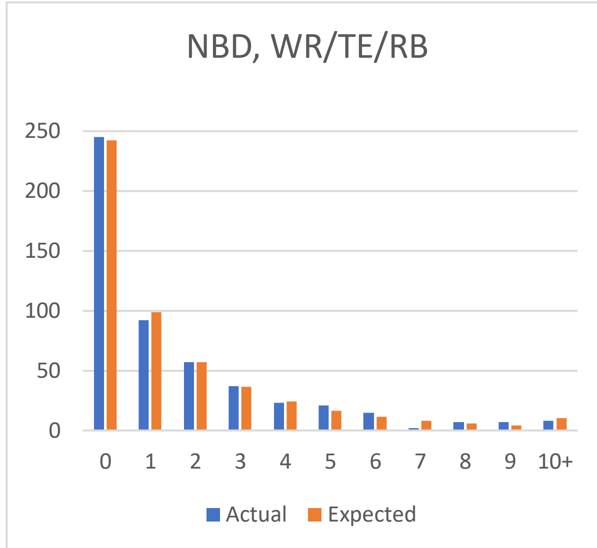
The consistency of these datasets is one concern, as I was unable to determine how players were chosen to be included for the different sites.

Fitting the basic NBD model

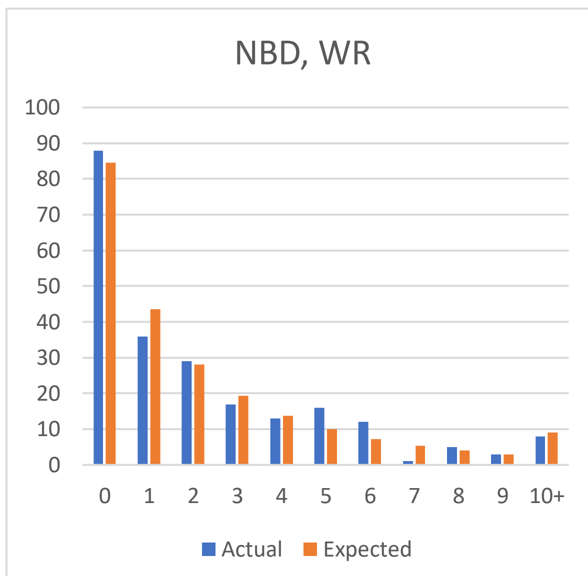
We begin by fitting an NBD model to the entire dataset (QB, WR, TE, and RB), finding parameters $r = 0.411$ and $\alpha = 0.294$. Among all offensive NFL players, there is high heterogeneity in receiving TD "propensity", logical as we are considering players in vastly different positions. Interestingly, the model fits decently upon inspection (p-value of 0.133).



We then fit an NBD model on only the positions of interest (WR, TE and RB), finding parameters $r = 0.542$ and $\alpha = 0.334$. As we expected, heterogeneity decreases (r increases) as we limit our population now to just pass-catchers. However, there remains high heterogeneity among NFL pass-catchers, and it does not appear to be the case that well-performing receivers are simply "lucky". The NBD model fit looks a bit better here, which is supported by a chi-square GOF test (p-value of 0.227).



Finally, we consider just running the model on just wide receivers, our most narrow subgroup. The resulting NBD model has parameters $r = 0.664$ and $\alpha = 0.290$, continuing the intuitive trend of heterogeneity decreasing as we narrow down the NFL population; the fit has seemingly gotten slightly worse, as shown by the histogram and GOF test (p-value of 0.129).

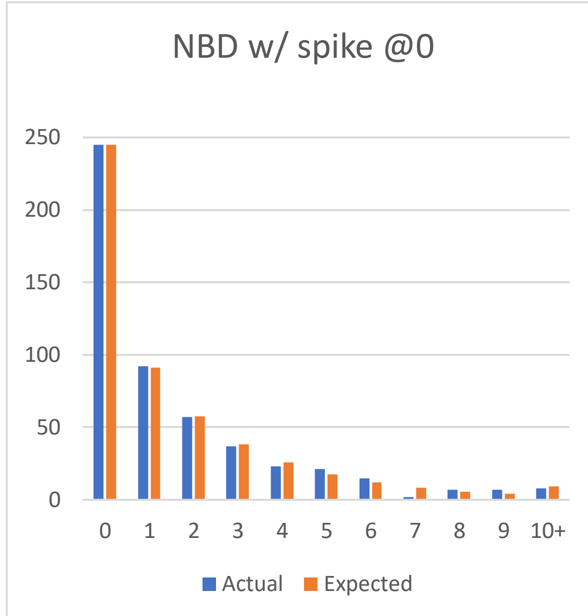


Based both on the fit of the model and our criteria decided above, **we proceed using the WR/TE/RB** dataset.

We note that in all three models, the NBD model greatly overpredicts the number of players that score 1 or 7 touchdowns.

More NBD models!

Next, upon noticing that there the model slightly predicts too few zero's and far too many one's, we fit a spike at zero in our NBD model. The resulting parameters are $r = 0.779$, $\alpha = 0.404$, and a spike proportion of 0.158. Right of the bat, the histogram seems to show a very strong fit! This is supported by the GOF test as well (p-value of 0.231); however, after long thought, I still decided to **leave out the spike**. Intuitively, one could say that there are some NFL players who never get the opportunity to truly "spin the Poisson wheel", but practically I could not justify where that line would be drawn. What defines the 15.8% of NFL players that are not part of our story? As I could not characterize this mystery group, I decided to move on despite the strong fit.



Next, to check the robustness of the fit, I fit the NBD model using summary statistics: first I used the means and zeros method, finding $r = 0.528, \alpha = 0.326$. Using the mean and variance, I similarly found that $r = 0.611, \alpha = 0.377$. Although these slightly deviated from the parameters of the basic model, to me this confirmed the robustness of the basic NBD model.

Finally, I fit zero-missing models, namely the shifted and truncated NBD models.

I got $r = 0.884, \alpha = 0.420$ for the shifted, and $r = 0.779, \alpha = 0.404$ for the truncated. For the truncated as well, the model suggested an effective zeros population of 164, much lower than the actual number of zero-TD scorers in the original dataset. Similar to what the spike suggested, this meant that only 164 receivers were part of our story, but as this did not make logical sense to me, I proceeded using the basic model.

	All-NFL	WR/TE/RB	WR	Spike (0.158)	Means/zeros	MoM	Shifted	Truncated
r	0.411	0.542	0.664	0.779	0.528	0.611	0.884	0.779
α	0.294	0.334	0.290	0.404	0.326	0.377	0.420	0.404

<u>Historical analysis</u>

After running the basic NBD model with the chosen positions on the 1987 and 2007 dataset, we find respective r and alpha parameters of 0.502 and 0.297, and then 0.510 and 0.485. This suggests that the heterogenous nature of receiving touchdown scoring has not changed significantly as the NFL has evolved.

<u>Further analysis</u>

In our study we treated $t = 1$ to mean a single NFL season, when another logical approach is to treat a single unit of time as one "opportunity", be it a target or a snap played. Given our findings that NFL receivers' propensities to score touchdowns are highly heterogenous, with opportunity baked into this result, running a model adjusted for individual opportunities could have implications about if there is any significant "skill" involved in catching touchdowns.