

Statistical Odds & Ends

What is Harrell's C-index?

Posted on [October 26, 2019](#)

Harrell's C-index (also known as the **concordance index**) introduced in Harrell et al. 1982, is a goodness of fit measure for models which produce risk scores. It is commonly used to evaluate risk models in survival analysis, where data may be censored.

Survival data

For concreteness, let's imagine that we are in the clinical setting. For a given patient, we are interested in predicting his/her "time-to-disease", i.e. the length of time until he/she develops the disease of interest, given covariates (e.g. demographic information) that we have on hand today. To train a model to make these predictions, we have n patients with their covariate information X_1, \dots, X_p and a "time-to-event" response T .

Survival analysis is different from typical regression/classification because of the nature of the response T . For a given disease, not all patients are going to develop the disease. In clinical settings, it's also possible that the patient never comes for a follow-up visit, so we won't know if the patient developed the disease or not. Hence, for the i th patient, the "time-to-event" response T_i is either

1. The actual time-to-disease if we get to observe it, or
2. The last time at which we know that the patient did not have the disease.

To know which of the two cases happened, we have an auxiliary variable d such that $d_i = 1$ if we got to see the disease, $d_i = 0$ otherwise. If $d_i = 0$, patient i 's time-to-event is said to be "censored".

Harrell's C-index

The intuition behind Harrell's C-index is as follows. For patient i , our risk model assigns a risk score η_i . If our risk model is any good, patients who had shorter times-to-disease should have higher risk scores. Boiling this intuition down to two patients: **the patient with the higher risk score should have a shorter time-to-disease.**

We can compute the C-index in the following way: For every pair of patients i and j (with $i \neq j$), look at their risk scores and times-to-event.

1. If both T_i and T_j are not censored, then we can observe when both patients got the disease. We say that the pair (i, j) is a **concordant pair** if $\eta_i > \eta_j$ and $T_i < T_j$, and it is a **discordant pair** if $\eta_i > \eta_j$ and $T_i > T_j$.
2. If both T_i and T_j are censored, then we don't know who got the disease first (if at all), so we don't consider this pair in the computation.
3. If one of T_i and T_j is censored, we only observe one disease. Let's say we observe patient i getting disease at time T_i , and that T_j is censored. (The same logic holds for the reverse situation.)
 - A. If $T_j < T_i$, then we don't know for sure who got the disease first, so we don't consider this pair in the computation.
 - B. If $T_j > T_i$, then we know for sure that patient i got the disease first. Hence, (i, j) is a **concordant pair** if $\eta_i > \eta_j$, and is a **discordant pair** if $\eta_i < \eta_j$.

Harrell's C-index is simply

$$c = \frac{\# \text{ concordant pairs}}{\# \text{ concordant pairs} + \# \text{ discordant pairs}}.$$

The logic above can be expressed succinctly in a formula (taken from Reference 2):

$$c = \frac{\sum_{i \neq j} 1\{\eta_i < \eta_j\} 1\{T_i > T_j\} d_j}{\sum_{i \neq j} 1\{T_i > T_j\} d_j}.$$

Values of c near 0.5 indicate that the risk score predictions are no better than a coin flip in determining which patient will live longer. Values near 1 indicate that the risk scores are good at determining which of two patients will have the disease first. Values near 0 means that the risk scores are worse than a coin flip: you might be better off concluding the opposite of what the risk scores tell you.

Harrell's C-index for continuous data

Of course, one can compute the C-index if none of the data is censored. In that case, all pairs such that $T_i \neq T_j$ will be included in the computation.

Harrell's C-index for binary data

As Reference 3 suggests, the concept of the C-index can be easily ported over to binary data. In this setting, a high risk score prediction means more likely to be 1 than a 0. We only consider pairs where subject i 's response is a 1 and subject j 's response is a 0. The pair is concordant if $\eta_i > \eta_j$, and discordant if $\eta_i < \eta_j$.

References:

1. Harrell Jr, F. E. et al. (1982). [Evaluating the yield of medical tests.](#)
2. Schmid, M. et al. (2016). [On the use of Harrell's C for clinical risk prediction via random survival forests.](#)
3. Statistics How To. [What is a C-Statistic?](#)

Advertisements



**When Neuropathy Hits, Try
This (It's Genius!)**

Sponsored by NerveShieldPlus

This entry was posted in [Uncategorized](#) and tagged [concordance](#), [harrell's c-index](#), [survival analysis](#) by [kjtay](#). Bookmark the [permalink](#) [\[https://statisticaloddsandends.wordpress.com/2019/10/26/what-is-harrells-c-index/\]](https://statisticaloddsandends.wordpress.com/2019/10/26/what-is-harrells-c-index/).

8 THOUGHTS ON “WHAT IS HARRELL’S C-INDEX?”



[kingcrimsontenseconds](#)

on [February 21, 2020 at 7:57 pm](#) said:

Thanks for the great explanation! +1337

★ Like



[nkasana](#)

on [March 2, 2021 at 9:47 pm](#) said:

Thanks very much for this, so helpful!

★ Like



[Joe Wu](#)

on [March 19, 2021 at 4:16 pm](#) said:

How do you implement it? Do you have a code (SAS or R code) to calculate it? How do you get the risk score ?

★ Like



[kjtay](#)

on [March 19, 2021 at 9:46 pm](#) said:

The risk scores for the individuals are whatever your risk model assigns. Harrell’s C-index is a way to evaluate how good the risk model is. In R, the glmnet package as a function called Cindex() you can use. It is a wrapper for the survival package’s concordance() function. (I can’t comment on SAS code as I am not a SAS user.)

★ Like



Mette Pedersen

on **February 16, 2022 at 9:22 am** said:

Thank you for a very useful explanation.

I hope you can help me with the following questions:

How close to one should the C-index be for the model to be considered a good fit?

If you have two (or more) models, how do you decide which have the best fit? When is a difference in C-index statistically significant? If a model has a C-index of 0.75 and another 0.76 I am guessing they have an equally good fit, but is a C-index of e.g. 0.78 statistically better than 0.75?

I hope you can help explain – thank you!

★ Like



kjoytay

on **February 20, 2022 at 5:01 am** said:

I believe there are statistical tests for testing differences in C-index. The compareC R package is one example: <https://cran.r-project.org/web/packages/compareC/index.html>

★ Like



Ian

on **April 19, 2023 at 1:10 pm** said:

Thanks for the clear explanation!

I think there's a typo at the end where you write the following:

"We only consider pairs where subject i's response is a 1 and subject j's response is a one. The pair is concordant if $\eta_i > \eta_j$, and discordant if $\eta_i < \eta_j$."

Shouldn't the "one" be "0" instead?

★ Like



kjaytay

on April 19, 2023 at 3:29 pm said:

Good catch Ian! We should only consider pairs where one subject has response 1 and the other has response 0. I've made the change in the post. Thanks!

★ Like