Photo-z comparison test

Metrics

This is a brief explanation on how the metrics are calculated for this comparison test, any suggestion or comment is very valuable. For each submission we have photo-z values and a error associated with it or a $p(z)$ distribution for each galaxy. We are using the weight files, on each sample we have a vector $\vec{\omega}$ of weights for each one of the N galaxy on each sample where $\sum \omega_i = 1$. If no weights are use, one can simply assign $\omega_i = 1./N$ to each galaxy. We define $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ and the following statistics, which are computed for the whole sample and for each bin in photometric redshift of width 0.1:

1. bias($< \Delta z >$) as:
$$\text{bias} = \frac{\sum \omega_i \Delta z_i}{\sum \omega_i} \tag{1}$$

2. σ as:
$$\sigma = \left(\frac{\sum \omega_i \left(\Delta z_i - \text{bias}\right)^2}{\sum \omega_i} \right)^{\frac{1}{2}} \tag{2}$$

3. median (Δz_{50}) as the median of the Δz distribution (or an approximation to it):
$$P_{50} = P(\Delta z \leq \text{median}) = \int_0^{\text{median}} \omega(\Delta z) d(\Delta z) = \frac{1}{2} \tag{3}$$

4. σ_{68} as the half of the width of the distribution, measured with respect to the median, where 68% of the data is enclosed. This is computed as:
$$\sigma_{68} = \frac{1}{2} \left(P_{84} - P_{16}\right) \tag{4}$$

5. frac($> 2\sigma$) as the fraction of outliers above the $2 - \sigma$ level computed as:
$$\text{frac}(> 2\sigma) = \frac{\sum W_i}{\sum \omega_i} \tag{5}$$
where,
$$W_i = \begin{cases} \omega_i, & \text{if } |\Delta z_i - \text{bias}| > 2\sigma \\ 0, & \text{if } |\Delta z_i - \text{bias}| \leq 2\sigma \end{cases}$$

6. frac($> 3\sigma$) as the fraction of outliers above the $3 - \sigma$ level computed as:
$$\text{frac}(> 3\sigma) = \frac{\sum W_i}{\sum \omega_i} \tag{6}$$
where,
$$W_i = \begin{cases} \omega_i, & \text{if } |\Delta z_i - \text{bias}| > 3\sigma \\ 0, & \text{if } |\Delta z_i - \text{bias}| \leq 3\sigma \end{cases}$$

7. $\text{bias}_{\text{norm}}$: Is the mean of the distribution of the normalized Δz by its error. In a unbiased case this distribution should resemble a normal distribution with mean 0 and sigma 1. We define $\Delta z' = \frac{z_{\text{phot}} - z_{\text{spec}}}{\epsilon_{\text{phot}}}$ where ϵ_{phot} is the computed error of the photometric redshift. Then:

$$\text{bias}_{\text{norm}} = \frac{\sum \omega_i \Delta z'_i}{\sum \omega_i} \tag{7}$$

8. σ_{norm} as[†]:

$$\sigma_{\text{norm}} = \left(\frac{\sum \omega_i \left(\Delta z'_i - \text{bias}_{\text{norm}} \right)^2}{\sum \omega_i} \right)^{\frac{1}{2}} \tag{8}$$

9. N_{poisson} : Is a measurement to quantify how close the distribution of $N(z_{\text{phot}})$ is to the distribution of $N(z_{\text{spec}})$, this is computed for each photometric redshift bin as the difference of $N(z_{\text{phot}}) - N(z_{\text{spec}})$ normalized by Poisson fluctuations, on each bin:

$$n_{\text{poisson}} = \frac{\left(\sum\limits_{z^i_{\text{phot}} \,\epsilon\, \text{bin}} \omega_i N - \sum\limits_{z^i_{\text{spec}} \,\epsilon\, \text{bin}} \omega_i N \right)}{\sqrt{\sum\limits_{z^i_{\text{spec}} \,\epsilon\, \text{bin}} \omega_i N}}$$

Then N_{poisson} is computed as the RMS of the previous quantity:

$$N_{\text{poisson}} = \left(\frac{1}{n_{bins}} \sum_{j=1}^{n_{bins}} n^2_{\text{poisson},j} \right)^{\frac{1}{2}} \tag{9}$$

10. KS: Is the Kolmogorov-Smirnov test that quantifies whether the two redshift distributions (photometric $N(z)$ and spectroscopic $N(z)$) are drawn from the same distribution, it provides one robust value to compare both distributions that does not depend on the binning. Can be thought as the maximum distance between both cumulative distributions. In this case the lower this maximum distance the better and both distribution are more similar. For both distributions the empirical distribution function is calculated as:

$$F_{\text{spec}}(z) = \frac{\sum\limits_{i=1}^{N} \Omega_{z^i_{\text{spec}} < z}}{\sum \omega_i}$$

where,

$$\Omega_{z^i_{\text{spec}} < z} = \begin{cases} \omega_i, & \text{if } z_{\text{spec},i} < z \\ 0, & \text{otherwise} \end{cases}$$

Similarly the empirical distribution function is computed for the values of z_{phot}, then the KS statistic is computed as:

$$\text{KS} = \max_z \left(||F_{\text{phot}}(z) - F_{\text{spec}}(z)|| \right) \tag{10}$$

For the submissions with $p(z)$ for each galaxy, these same metrics are computed but taking into account the $p(z)$ distribution for each galaxy on each case (which takes much longer to calculate), more details can be found within the code.

[†]Note that $\text{bias}_{\text{norm}}$ and σ_{norm} are different from those values shown in the bottom-left panel of Figure 1 produced by the plotting routine, those values are the result of a Gaussian fit to $\Delta z'$ where only values of $\Delta z' < 5$ are considered for the fit for illustration purposes