How Much Is an S-Unit?

A statistical method to evaluate amateur radio communication systems

Jürgen A. Weigl OE5CWL/OE6CWL
weigl+info@magnet.at

Have you ever wondered what the difference is between the vertical at your station and that Tribander at the station of your good buddy around the corner? Or are you going to operate the next contest as a guest operator and would like to know how competitive this station will be? Analyzing the logbook with statistical methods may be one way to find the answer to some of your questions.

Although statistical methods are used today in many fields, we do not recall any publication in the amateur radio literature how these methods can be successfully used by hams. The author used statistical methods first some years ago, when trying to find out how a new 40 m antenna performed compared to the old one, which was already removed. Since then a lot of time was spent to find out how statistics could be used by amateurs to ensure reasonable results.

The Signal Report:

When you are in contact with a new station, especially if you are chasing DX, you are always anxious on knowing how strong your signal is. For this purpose we amateurs use the RST-code. In other short-wave communication systems different codes are used. For example Short-wave Broadcast DXers use a code called the SINPO- (or SINFO-) code for reception reports. SINPO means S for signal strength, I for Interference, N for Noise, P (or F) for fading and O for overall performance. Each value is specified by a figure between 1 and 5 (Table 1). As you can see this code has some more information than our RST-system. The reception reports are used by technicians of many broadcast stations to find out the necessary changes to improve the signal. We will use similar methods here with the signal reports in our RST-code.

S - QSA Signal strength		**I - QRM** Interference		**N - QRN** Noise		**P - QSB** Fading		**O - QRK** Overall merit	
5	excellent	5	nil	5	nil	5	nil	5	excellent
4	good	4	slight	4	slight	4	slight	4	good
3	fair	3	moderate	3	moderate	3	moderate	3	fair
2	poor	2	severe	2	severe	2	severe	2	poor
1	barely audible	1	extreme	1	extreme	1	extreme	1	unusable

Table 1: SINPO-code used for reception reports for broadcast stations

Our RST-code consists of two figures on phone or three figures in CW. In this code R stands for readability, S for signal strength and T for tone. While there are 5 possible figures (1 to 5) for the readability, in most cases we will get a 5 report, meaning that readability of our signal is perfect. Although readability is by far not always perfect, amateurs tend to let the R be a 5. Therefore this "variable" is of not much value for our investigation. T is only used in CW and with the solid-state transceivers we use today it really is (or has to be) in almost all cases a 9, meaning that our tone is perfect. Therefore this also has no value for our approach. Somewhat different is the S in the exchanged reports. This value varies between S1 and S9plus. With the figure for S you receive in the report during your QSO the other station will rate your signal. You have to take in mind, that your signal will be rated, but in almost all cases you will not receive a measured value for your signal. It is important for our investigation to consider the S-unit not purely as a technical value. This has several reasons:

a.) even in our modern transceivers the S-meters are not very accurate. The S-meter reading can be far off the real value. There also can be some variance when changing bands.

b.) many amateurs just estimate the signal strength by ear, S-units corresponding to definitions like the one given in table 2.

c.) the report will be psychologically influenced and a rare station may perhaps be more likely to receive a S9

Therefore we have to consider the S-report as some rating for our signal, influenced by many factors not under our control. Nevertheless we will see, when using a sample big enough we will be able to minimize these effects and get reasonable results.

S - Signal strength	
1	faint signals, barely perceptible
2	very weak signals
3	weak signals
4	fair signals
5	fairly good signals
6	good signals
7	moderately strong signals
8	strong signals
9	extremely strong signals

Table 2: Definitions for S-units [reference 1]

For our investigation all statements we will give refer to S-units as commonly used by short-wave amateurs. We will not speak of a specific voltage at the receiving point. We will treat S-reports more like an opinion poll, with a question like "how much does the other station like my signal?" Speaking of opinion polls, it is obvious that these instruments today are indispensable in politics and marketing. While asking one single person, for whom he will vote in the next elections, will not really give you much information about the prospected results, asking the question a reasonable sample of people, will give you the necessary data to optimize your campaign. And we will treat S-reports just that way, some sort of opinion poll, how much the other station likes our signal. And the operator at the other station can choose from a scale from 1 to

9plus. Now you may ask, if this makes any sense. Well, on the one hand you will see, that using such a method you get clear indications for technical improvements. On the other hand, isn't a contest also somewhat like an election or some marketing campaign? Instead of the question: "whom are you going to elect?" or "what are you going to buy?" you just have the question "whom are you going to give a call in the contest?" And it's obvious for every contester that on a scale, measuring the probability to get that call, you'd like to be as far up as possible.

When we are speaking about contesting, there is one point we have to keep in mind. Although a contest would be the best situation to make such an analysis, as we work thousands of other stations within a very short period of time, we can not use contest reports for our study. The reason is obvious: the exchanged report is in almost all cases 59, making the report by itself meaningless. Therefore we must not use contest reports for our study, but have to use everyday QSO outside of any contest for our investigation.

Statistical Treatment of Signal-Reports:

We will start with an examination of the typical distribution of signal reports. The author has been chief operator of an Austrian club station and has therefore access to a great amount of data of short-wave QSOs. Let's first take a look on the club stations' reports on 20 m working Europe. When investigating data like this, it is essential to minimize any influence that would introduce possible errors in our evaluation. What are the possible influences for our example? First of all the output power and the antenna we use at the transmitting side. We can control this and do our investigation only for contacts with the same power level and the same antenna. In our example all contacts were made with an output power of 100 watt and the antenna was a simple trap-vertical (14 AVQ from HyGain) about 35 meters from ground.

The second influence comes from the distance over which we make the contact. It's obvious, that our signal will be different when working DX or contacts on our own continent. In our example we try to control this influence be limiting our investigated contacts only to contacts with other European stations. The next influence comes from the band conditions: while conditions in general terms vary within the period of a solar cycle, they also vary on an everyday basis. We will control the first influence, meaning, that we consider only contacts within a similar period of solar activity, described by the relative and smoothed number of sunspots. We will not try to control the everyday variation of the ionospheric conditions, but we will see, that with a sample big enough, this influence will be minimized. There is one other influence that we cannot control: the receiver, operator and antenna (gain) at the other end of the QSO. But we will assume, that when we use a sample big enough, we will get a representative profile of possible stations within our target area.

Now we searched the log of the club-station within a period of one year for 20 m-contacts with European stations. There were two different operators who made such contacts. Operator 1 made 115 contacts, operator 2 had 106 contacts. To avoid any possible influence from the operator himself, we investigated every operator's contacts alone. Table 3 shows the number of S-reports for each S-unit and each operator. With this list, we are now able to compute the average signal report. This is done by multiplying the number of reports of one category with their value, add it up and divide the result by the total number of contacts. This can be mathematically described by equation 1:

$$\bar{s} = \frac{1}{n} \sum_{i=1}^{10} n_i * s_i \qquad \text{(cquation 1)}$$

$\bar{s}$ ... average signal report
n ..... number of contacts
n_i ... number of reports with value s_i
s_i ... value of signal report ($i=1$, $s_i = 1$; $i=9$, $s_i = 9$)

The value of each category is the corresponding S-value. We now have one problem: what is the value of a report 20 dB over S9? At first we do not take into account how much over S9 a report might be. We simply put all over S9 reports into the category 9plus. It is then reasonable, using a linear scale, to give the 9plus category a value of 10.

	S-Report	9+	9	8	7	6	5	4	3	2	1	Total
Op 1	**No. Of Reports**	11	35	26	13	15	11	4	0	0	0	115
	Percentage	9.57	30.43	22.61	11.30	13.04	9.57	3.48	0	0	0	100
	CDF	9.57	40	62.61	73.91	86.96	96.52	100	100	100	100	
Op 2	**No. Of Reports**	11	35	21	12	15	7	2	3	0	0	106
	Percentage	10.38	33.02	19.81	11.32	14.15	6.60	1.89	2.83	0	0	100
	CDF	10.38	43.40	63.21	74.53	88.68	95.28	97.17	100	100	100	

Table 3: Signal reports for an Austrian Club Station (OE6XRG) for contacts on 20 m with Europe. Two different operators during the same time period.

Computing the average signal report for both operators in our example brings an average report of 7.696 for operator one and 7.726 for operator two. Although there were two different operators, working completely different stations within Europe and operating at completely different times and days within the same time period, the average signal report for them differs by only 0.03 S-unit!

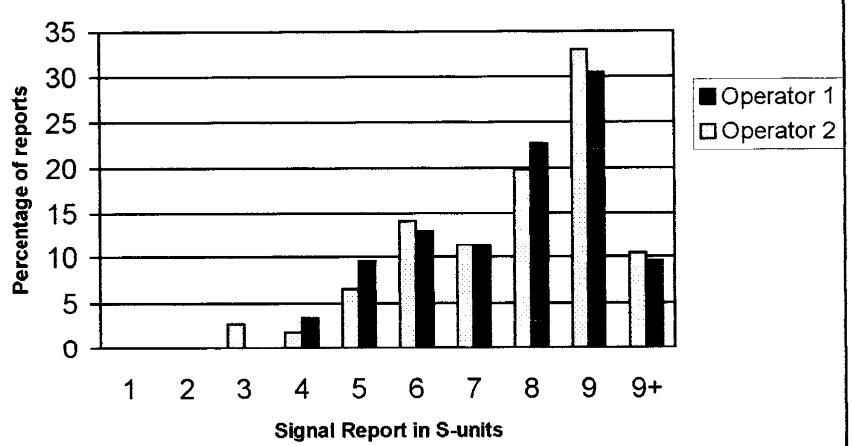
But there is also another way to treat the received reports, which will give us even more information about our signals. For this we take a look at the distribution of received reports. We start calculating the percentage of signal reports for each category. This is described by equation 2. Operator one had 35 S9 reports by a total of 115 contacts, so there are 30.44 % (= 35/115*100%) S9 reports. This is done for all other reports and the other operator.

$$P_{[\%]} = \frac{100}{n} n_i \qquad (2)$$

$P_{[\%]}$ ... Percentage of signal report with the value s_i
n .... number of observed contacts
n_i ... number of contacts with the value s_i
s_i ... value of signal report ($i=1$, $s_i = 1$; $i=9$, $s_i = 9$)

The results are shown in Table 3 and Figure 1 where we see the percentage of reports for each value of reports and for both operators.

Figure 1
Percentage of signal reports for club station OE6XRG, 20m to Europe



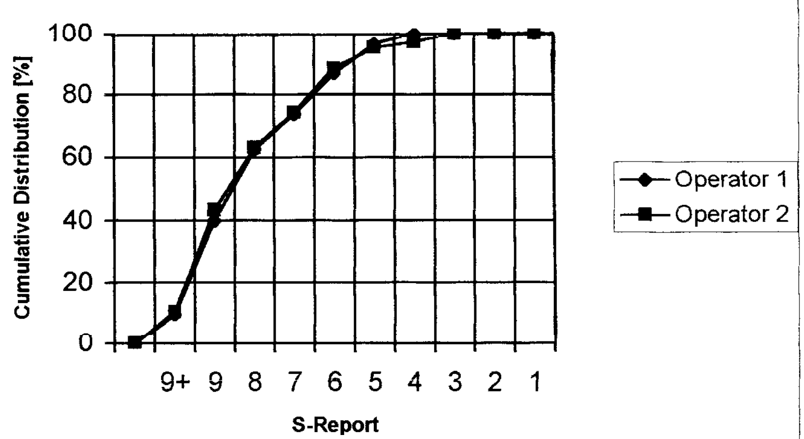
Although we see from this diagram that at about 30 % of the time we receive reports with an S9, there is another, more interesting way to look at the distribution of reports. This is the **cumulative-distribution-function** (CDF): it is convenient to indicate for each value the percentage of reports having values better than or equal to that value. This is described by equation 3. The S9plus percentage for operator one is 9.57% and for S9 30.43%, therefore the cumulative distribution for 9plus is 9.57% and for S9 40% (30.43 + 9.57), which means that 40 % of the reports are S9 or better.

$$C_m = \sum_{i=m}^{10} \frac{100}{n} n_i \qquad (3)$$

C_m ... Cumulative distribution of signal reports for the value s_m [%]
n .... number of observed contacts
n_i ... number of contacts with the value s_i
m ... ordinal number of signal report (m=1, $s_m = 1$; m=9, $s_m = 9$; m=10, $s_m = 9^+$)

It is convenient to show this distribution graphically as done by Figure 2.

Figure 2
Cumulative distribution function for club station OE6XRG, 20m to Europe



This distribution is typical for most of the cases the author has analyzed, although the curve may be moved more to the left or right side of the diagram. Please note, that the value at 50 % is not identical with the average signal report! Figure 2 clearly indicates how close the results are for both operators, concerning the cumulative distribution of signal reports. While taking one single report or even a handful of reports does not really give you much information about the performance of your station, doing an analysis as this one with a reasonable sample will bring you some interesting results. From the curve in Figure 2 (or from Table 3) we see, that on 20 m, working Europe this station will get an S9 or better signal report in 40 % of the cases, and we will get an S8 or better report in about 60 % of the cases. We can also put it in other words: the probability to be rated S9 or better by another European station is only 40 %. Looking at this result as a contester we can say, that this station does not seem to be very attractive. The author participated in some contests from this club station and it really was not very competitive. But what else would you expect from 100 Watts and a multiband vertical, even if that vertical is high up in the sky?

But it is important to see, that for the cumulative distribution function the same happened as for the average signal report. The difference between both operators for the average signal report was only 0.03 S-units. The cumulative distribution function is almost identical for both operators, although the operators worked completely different stations within Europe and were on the air at completely different times and days, but within the same time period of one year. Therefore we can conclude, that we really get reasonable results, by analyzing a bit more than 100 contacts.

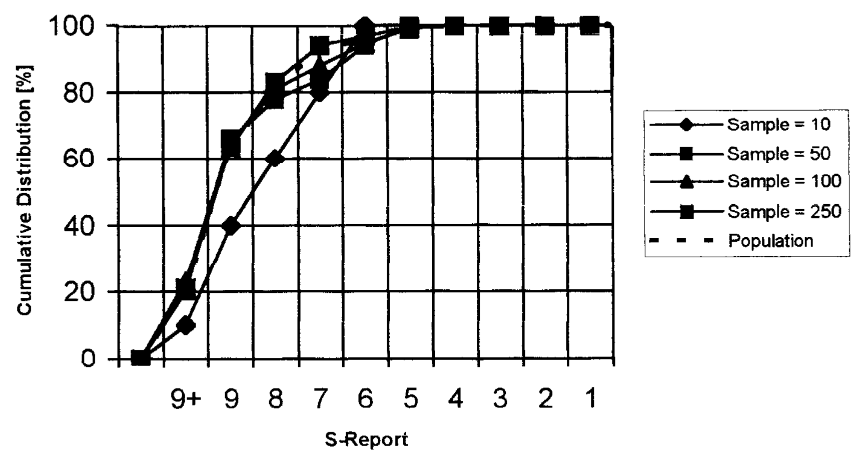
Before we go on with some examples how to use the CDF for comparing different stations or equipment, let us take a look how results may differ when we change the size of our sample. A sample is defined in statistics as a subclass of a population. In our case the population was all possible contacts with European stations on 20 m, the sample of this population was all contacts operator one or operator two made during the observed time period. The population could as well be all possible contacts with Japan - a sample could be the last 50 Japanese contacts you made. To get reasonable results we have to choose a sample big enough to get a low error. As we have seen from our samples of 115 contacts from operator one and 106 contacts from operator two, the results are very close, which suggests, that these samples were big enough.

A sample can range from one contact to any amount of possible contacts. It's obvious that with one contact we can not get any realistic statistic inference, while when the number of contacts considered is increased, we can be pretty sure of our average signal report and the report distribution. To make this more clear, we will bring an example: we have worked 500 European stations on 80. We consider these 500 European stations to be our population and we can calculate the average signal report and the signal report distribution. Now we can program a computer or use some other methods, to produce random signal reports representing that distribution. We choose a sample size and compare the results with the result we have for our entire population (500 European contacts on 80). Figure 3 and Table 4 show, that with increasing the sample size, we get closer to reality. You can see that the results get very accurate when you have a sample big enough. This is done by simply random sampling, where we have two conditions necessary to speak of a random sample:

a.) each member of the population is just as likely to be included in the sample as any other member

b.) the likelihood that any given member of the population will be included in the sample is affected equally by the inclusion of any other particular member

Figure 3
Cumulative distribution function for different sample sizes



Sample	9+	9	8	7	6	5	4	3	2	1	Average Signal report
10	1	3	2	2	2	0	0	0	0	0	7.90
50	10	23	6	3	5	3	0	0	0	0	8.42
100	23	42	16	7	7	4	1	0	0	0	8.51
200	42	85	39	22	6	4	2	0	0	0	8.575
500	104	215	91	51	22	14	3	0	0	0	8.548
CDF 10	10	40	60	80	100	100	100	100	100	100	
CDF 50	20	66	78	84	94	100	100	100	100	100	
CDF 100	23	65	81	88	95	99	100	100	100	100	
CDF 200	21	63.5	83	94	97	99	100	100	100	100	
CDF 500	20.8	63.8	82	92.2	96.6	99.4	100	100	100	100	

Table 4: Number of reports, average signal report and cumulative distribution for different sample sizes

After investigating a sample of contacts, we can state statistical inference. Of course we have to point out, that we are working with statistic and our results are accurate to a certain level of confidence. This is called the confidence interval and represents the probability that our result will be accurate within a certain level of error. Also this sounds complicated, a simple example may help: Just remember what happens when you flip a coin. Each flip is classified to heads or tails. The probability for heads is as high as for tails. There are two possible sides, so we can state the probability to be 1:2 or for calculating 0.5 (50%). The definition for probability of obtaining a specific value out of a population is the proportion of the population having that value. In our example of tossing a coin, the population has two members, the proportion of heads and tails is equal. It's obvious that when you have a high number of tosses there will almost be equal amounts of tails and heads. The higher the number of tosses, the more will the percentage of heads and tails approximate the 50% mark. Just imagine you do 10 or 100 tosses. How high is the probability for the maximum possible error? The error would be maximum within our observation, if all 10 or 100 tosses bring the same result. Now high is the probability, that in a series of 10 or 100 tosses you only have tails? You will agree that the probability for such an event is very low. For such a simple example it is very easy to calculate the probability: for 10 tosses you have a probability of about 0.1 %, for 100 tosses the probability is only about $8 * 10^{-29}$ %!

Exactly the same is happening with our distribution of signal reports. Even with a small sample the probability for the worst case, that means for getting the maximum possible error is very low and decreases as the number of observations, our sample size, is increased. For example taking our population of 500 European contacts, you may find that receiving a report of S 9plus is quite likely to happen, as there are 104 reports with that value out of total of 500. Now assume you only make 5 contacts. How high is the probability that you get a S9plus report for all 5 contacts? You can calculate this and find out that the probability for such an event is only 0.036%!

For our purpose taking a sample size of, say, 100 contacts we can be pretty sure, better say confident (as we are working at a specific confidence level) that we get reasonable results. Take a look at Figure 3: even with a small sample size of only 10 to 20 contacts, you will already see the trend, while the results by itself will not yet be very accurate. Using both values, the average signal report as well as the cumulative distribution, is advisable. From both you will be able to state your statistical inference. We said, that we are treating S-Meter reports not so much as a technical value, but as some kind of opinion-poll. Then it may be interesting, what the size of a sample in an opinion poll is. The author does not know the values for opinion-polls in the USA, but here in Austria we have a population of about 8 million people, from whom about 5 million are adults. The typical size of a sample in our opinion-polls are between 500 and 2000 persons. That is only 0.01 to 0.04 percent of the population! This shows us that with about 100 contacts for our sample, we have chosen a rather big sample compared to the entire population.

Some words of caution:

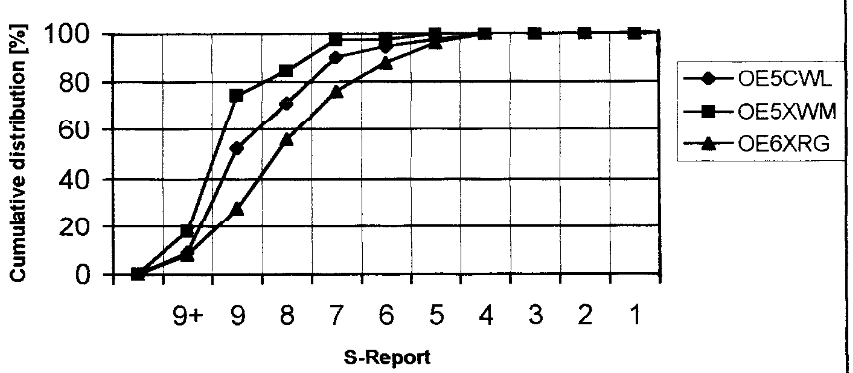
Before bringing some examples for practical applications some words of caution: when we use the signal report for investigation, we are always evaluating the entire communication system. That includes the equipment, the antenna as well as the operator at the station or the propagation conditions. So you have to make sure, that only the one thing in the whole system, which you want to evaluate, changes. For example, if you add an amplifier to your station and would like to find out how your signal improved, you should gather reports on one specific band with one specific target area and within a period without dramatic change in band conditions. Comparing results of contacts in the low part of the sunspot cycle with that in the maximum will be interesting, but will not bring you reasonable information how the amplifier performs.

Also your target area, which we called above the population (all possible contacts with Japan), should be clearly defined ahead. You might only take contacts within a 500 mile distance into account, or limit it to all DX-contacts or only to Russian stations etc. Of course the sample should be as big as possible as we have seen above. But now lets take a look at some possible applications of our statistical method:

Some examples

Using this method, we can easily find out, how different stations perform. At the same time the author has been chief operator at OE6XRG he owned a more competitive station together with his brother (OE5CUL). The setup there was 100 watt output into a tribander, which had 3 working elements on 20 and 5 elements on 15 and 10. This antenna was about 17 meters high. Unfortunately we don't have this station anymore, and the author now usually operates from another clubstation (OE5XWM) but using his own callsign (OE5CWL/5). This station has a three element tribander about 15 meters high and an output of 400 watt (which is the legal limit in Austria). Now it may be interesting to see, how these stations compare. Figure 4 shows the cumulative distribution of signal reports for all three stations for 20 m working North America. The contacts for this analysis were not within the same time period, but it was made sure, that the analyzed time periods had similar sunspot activity.

Figure 4
Cumulative distribution functions for three different stations working NA
OE5CWL: 100w to 3 el tribander at 17m
OE5XWM: 400w to 3 el tribander at 15m
OE6XRG: 100w to multiband trap vertical at 35m

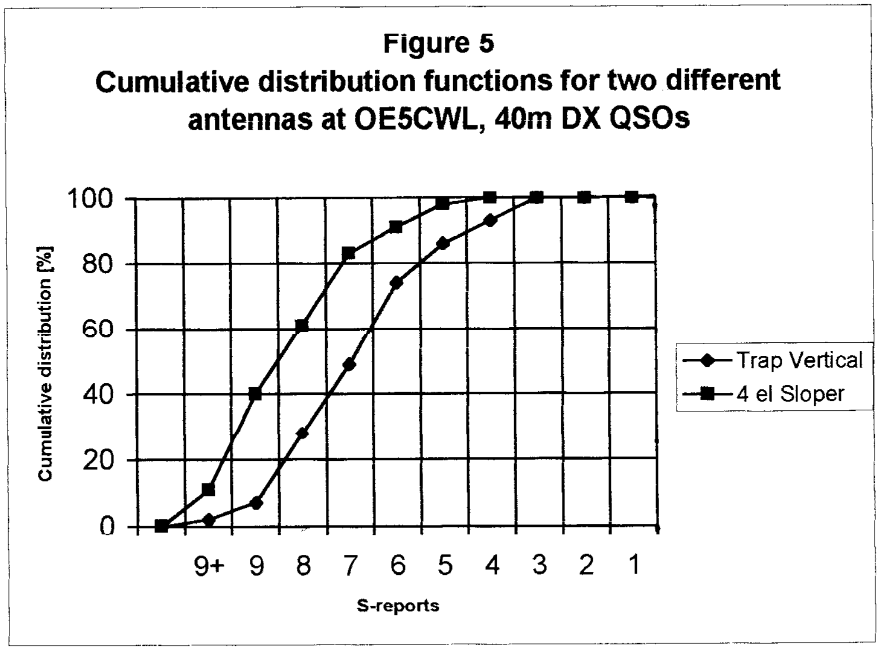


When we take a look at Figure 4 let's first compare the two stations with the same power level OE6XRG and OE5CWL. The average signal report differed by 0.887 S-units in favor of the Yagi at OE5CWL. At the 50 % level, which the author suggests to use as a standard (but remember this is not the same as the average signal report) the difference is about one S-unit. While it was possible to keep a frequency and work stations by calling CQ in contests at OE5CWL, this was not possible at the OE6XRG-location. Here only S&P produced a reasonable rate. Another interpretation of the results makes this clear: as Figure 4 indicates the probability to receive a report S9 or better in North-America is twice as high with the Yagi at OE5CWL than with the vertical at OE6XRG. We think this is a very important conclusion for contesting. Although the difference between both stations is only roughly one S-unit, the chance to have a good signal in North America has doubled! Keep this in mind, when optimizing your contest-station. There is a part of the CDF that is rather steep and with every improvement you will move the CDF more to the left. In this steep part of the curve you get a dramatic improvement, while it will not make too much difference at the flat upper and lower end of the curve.

Now lets take a look at the results of the third station OE5XWM. The Yagi at OE5XWM is not the same type as at OE5CWL. The later one is three elements on 20 with a boom length of 7.5 meters and is a bit higher than the Yagi at OE5XWM that uses a boom length of 5 meters. We think that this Yagi is somewhat inferior to the Yagi used at OE5CWL, but the higher output at OE5XWM is clearly indicated by the cumulative distribution function. The probability for a S9 or better signal in North America is now about 75 %, which means an improvement of about 50 % compared to the signal of OE5CWL, who has a probability of about 50 %. Compared to OE6XRG the probability for a S9 or better signal is now almost three times higher!

With this method you can not only compare different stations, but also find out the results of improvements. As a matter of fact, the author developed this method, when he tried to find out, how two antennas compared. Some time ago our 40 m antenna was changed. We removed the old 14 AVQ Trap-Vertical and built up a 4 element sloping dipole array. The vertical was mounted on a flat metal roof. So it had a very good counterpoise and we had real good success with it working DX on 40. Nevertheless a directive antenna was always in our mind. Therefore we constructed a 4 element sloper system similar to the one described in reference 2. It used shortened dipoles and by a switching system one element was chosen as the driven element, while the remaining elements worked as reflectors.

This antenna really showed excellent results. Unfortunately it was impossible to compare both antennas on the air as the vertical had to be removed. Therefore we analyzed all 40 m DX-contacts within one year (before/after changing the antenna). Figure 5 shows the results and this really confirms the improvement with the new antenna.



The difference of both antennas at the 50 % mark is about 1.6 S-units, the average signal is about 1.4 S-units better. But again, what is much more important can be found at the S9-mark. The probability for a S9 or better report is now 5 times higher! The author thinks this observation is of more importance, than the difference in the average report. What really counts for the DXer or contester is to improve the probability to break through the pile-up. It's obvious that the increased probability for a strong signal in DX, now being five times higher, is much more worth than the average improvement by a bit more than one S-unit. From this point of view you may understand why any improvement, even less than the usual 3 dB, is so important when working DX or contests. Every fraction of a dB will move the CDF to the left and therefore increase the probability for you to be heard.

Of course there are more possible applications for our statistical method. Besides comparing different station or antennas, you may evaluate how your signal improved with that new amplifier, or even to find out the improvement by your speech processor, where it's difficult if not impossible to give any dB-figure. The author hopes, that the described statistical methods will be useful for many amateurs and would appreciate any comments or criticism.

References:

[1] The R-S-T System, The Radio Amateur's Handbook, Chapter 24, Operating a Station, p647, ARRL, Newington, CT, 1976

[2] 7 MHz Sloper System, The ARRL Antenna Book, 18th Edition, Chapter 6, "Low-Frequency Antennas", p6-30, ARRL, Newington, CT, 1997