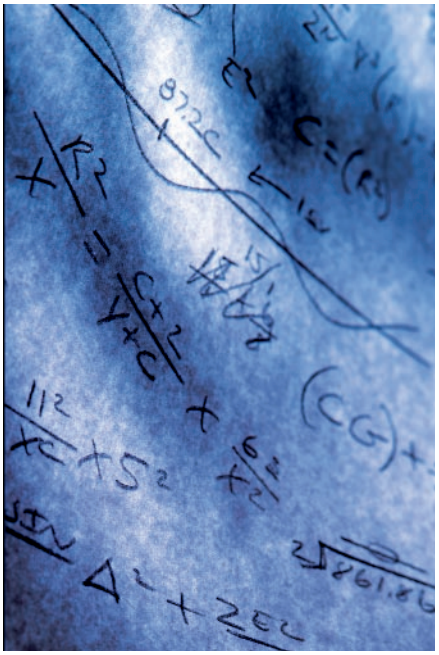

Statistical Tools for Setting In-Process Acceptance Criteria

Robert J. Seely, Louis Munyakazi, and John Haury

Regulatory approval and successful manufacturing depend on establishing meaningful and reasonable acceptance criteria for process validations and ongoing monitoring. Three methods are presented here to correct for the likely underestimation of process limits due to small samples.

Statistical approaches to in-process acceptance criteria are justified in light of FDA's emphasis on such approaches to setting specifications and establishing scientifically justifiable limits. CFR 211.110b, *Sampling and Testing of In-Process Materials and Drug Products*, says that "Valid in-process specifications for such characteristics shall be consistent with drug product final specifications and shall be derived from previous acceptable process average and process variability estimates where possible and determined by the application of suitable statistical procedures where appropriate" (1).

The difficulty of setting acceptance criteria depends on the availability and quality of historical data. Biotechnology products are often developed and introduced based on a limited number of production runs, and the manufacturing process may have changed during clinical development. Thus data may be insufficient for reliable estimation of process variance — and underestimation of process variance is common with small data sets. To correct that underestimation, we identify three statistical tools that allow the setting of acceptance criteria with ranges slightly wider than the conventional "mean ± 3 standard deviations," which is frequently the industry standard.

Acceptance limits determined using these statistical tools can be used for process validation and for continued process monitoring, demonstrating a state of statistical process control. After 15 or preferably 30 data points are obtained, a control chart based on the moving range is the best tool for monitoring process stability (2).

Materials

The example data used here are from a large-scale manufacturing process. The process has been validated, and the Biologics License Application (BLA) has been submitted.

For the chromatography step, we use a 1.6 meter diameter column with a bed volume of 500 L. The load material is an *E. coli* homogenate that is partially clarified by centrifugation and filtration. The lysate volume is such that the column cycles seven to eight times to work through the material. A full cleaning regime is run following each protein cycle. Following approval of the license, the process will be run at maximum capacity: The columns will run continuously and will not be placed in storage for the entire year of production. They will therefore run more than 700 cycles. Monitoring data, some of which are discussed, are used to set in-process acceptance criteria for validating the cycle

Table 1. Column yield example with 10 observations (81, 66, 93, 84, 84, 84, 97, 95, 92, 88), therefore nine degrees-of-freedom

	Textbook Method	Alternative 1[a]	Alternative 2[b]	Alternative 3[c]
Multiple of SD[d]	3	3.321	1.653t	3.197
UCL	113.4	116.2	119.8	115.1
LCL	59.4	56.6	53.0	57.7

[a] two-sided 95% prediction interval for *m* future runs based on *n* past runs
[b] one-sided 95% chi-square factor and two-sided 95% t-value (t=2.262)
[c] two-sided 95% tolerance interval controlling each tail to $p=0.05$
[d] standard deviation; mean equals 86.4; SD equals 8.98

Corresponding author ***Robert J. Seely*** *is a corporate quality technology scientist at Amgen, Inc., 3200 Walnut Street, Boulder, CO 80301, 303.401.1586, fax 303.938.6222, rseely@amgen.com.* ***Louis Munyakazi*** *is a project biostatistician, and* ***John Haury*** *is a senior manager, quality engineering, in the quality technologies group, Amgen, One Amgen Center Drive, Thousand Oaks, CA 91320-1789.*

number at scale, concurrent with commercial production.

Setting Acceptance Criteria

Acceptance criteria are selected using various general approaches in inspections and applications:

- Picking a range based on historical process behavior. The range must not be excessively large, or it will be difficult to defend.
- Assigning the previous high and low as the range. That is also acceptable but may not encompass all normal process variations.
- Using the limit of detection (LOD) for clearance studies. This approach is not recommended because of the increasing sensitivity of assays. Also, the LOD is often extremely low and beyond the scope of what is "reasonable and achievable."
- Basing the limits on statistics. This is the preferred method.
- Specifying no range, but looking for internal consistency. This approach can apply to secondary performance parameters such as pool volume or step time, where the mean can shift as the process is scaled or transferred, but consistency among the three to five validation runs is expected.

We recommend that limits be set based either on historical data or statistics when applied to primary performance parameters such as step yield and purity. Such

Table 2. Standard deviation multiplier table; distribution two-sided 100 $(1-\alpha)$% prediction intervals for m future observations using the results of the previous sample of n observations

$1-\alpha$	n	μ 1	2	3	4	5	6	7	8	9	10	12	16	20	40	60	80	100
0.9	4	2.631	3.329	3.742	4.032	4.255	4.435	4.585	4.713	4.826	4.925	5.096	5.357	5.555	6.143	6.468	6.692	6.861
	5	2.335	2.909	3.246	3.483	3.665	3.812	3.936	4.041	4.134	4.216	4.356	4.573	4.738	5.229	5.502	5.690	5.833
	6	2.177	2.685	2.982	3.190	3.350	3.480	3.589	3.682	3.764	3.836	3.060	4.153	4.299	4.736	4.980	5.148	5.276
	7	2.077	2.546	2.818	3.608	3.155	3.273	3.373	3.458	3.533	3.599	3.713	3.889	4.023	4.426	4.652	4.807	4.925
	8	2.010	2.452	2.706	2.884	3.021	3.132	3.225	3.305	3.375	3.437	3.543	3.708	3.834	4.213	4.425	4.571	4.683
	9	1.960	2.383	2.625	2.794	2.924	3.029	3.118	3.193	3.259	3.318	3.419	3.576	3.696	4.056	4.258	4.398	4.505
	10	1.923	2.331	5.564	2.726	2.851	2.951	3.036	3.108	3.172	3.228	3.325	3.476	3.591	3.936	4.131	4.265	4.368
	15	1.819	2.188	2.395	2.539	2.648	2.737	2.811	2.875	2.930	2.980	3.065	3.107	3.297	3.601	3.772	3.891	3.982
	20	1.772	2.122	2.318	2.454	2.556	2.630	2.709	2.768	2.820	2.866	2.945	3.068	3.162	3.445	3.604	3.715	3.800
	25	1.745	2.085	2.275	2.405	2.504	2.583	2.650	2.707	2.757	2.801	2.877	2.994	3.084	3.354	3.506	3.612	3.693
	30	1.727	2.061	2.246	2.373	2.470	2.547	2.612	2.667	2.716	2.758	2.832	2.846	3.033	3.204	3.442	3.544	3.623
	60	1.685	2.003	2.178	2.298	2.388	2.460	2.520	2.572	2.617	2.656	2.724	2.829	2.900	3.140	3.283	3.377	3.448
	120	1.665	1.976	2.146	2.262	2.349	2.419	2.477	2.526	2.569	2.607	2.673	2.773	2.849	3.078	3.206	3.294	3.362
	∞	1.645	1.049	2.114	2.226	2.311	2.378	2.434	2.481	2.523	2.560	2.622	2.718	2.791	3.008	3.129	3.212	3.276
0.95	4	3.558	4.412	4.023	5.285	5.564	5.789	5.978	6.140	6.282	6.407	6.622	6.054	7.206	7.954	8.370	8.655	8.872
	5	3.041	3.697	4.087	4.364	4.577	4.751	4.896	5.021	5.131	5.229	5.395	5.655	5.852	6.441	6.771	6.997	7.170
	6	2.777	3.333	3.662	3.896	4.076	4.223	4.346	4.452	4.545	4.626	4.770	4.901	5.150	5.665	5.949	6.145	6.295
	7	2.616	3.114	3.407	3.614	3.774	3.905	4.014	4.108	4.191	4.265	4.391	4.589	4.739	5.194	5.449	5.626	5.761
	8	2.508	2.968	3.236	3.426	3.573	3.692	3.792	3.879	3.954	4.022	4.138	4.319	4.457	4.876	5.112	5.275	5.400
	9	2.431	2.863	3.115	3.202	3.429	3.540	3.634	3.714	3.785	3.848	3.056	4.125	4.255	4.647	4.869	5.022	5.140
	10	2.373	2.785	3.024	3.192	3.321	3.426	3.515	3.591	3.658	3.717	3.820	3.980	4.162	4.474	4.685	4.831	4.942
	15	2.215	2.574	2.778	2.921	3.031	3.120	3.194	3.258	3.314	3.365	3.451	3.585	3.689	4.002	4.180	4.305	4.400
	20	2.145	2.480	2.670	2.801	2.902	2.983	3.052	3.110	3.162	3.208	3.286	3.409	3.503	3.788	3.951	4.064	4.151
	25	2.155	2.427	2.608	2.734	2.829	3.907	2.971	3.027	3.075	3.119	3.193	3.309	3.397	3.666	3.820	3.929	4.008
	30	2.079	2.393	2.569	2.690	2.783	2.857	2.920	2.973	3.020	3.062	3.133	3.244	3.320	3.587	3.734	3.837	3.915
	60	2.018	3.312	2.475	2.587	2.672	2.740	2.797	2.846	2.888	2.926	2.991	3.091	3.167	3.398	3.529	3.619	3.689
	120	1.988	2.274	2.431	2.538	2.619	2.685	2.739	2.785	2.826	2.862	2.923	3.018	3.090	3.308	3.430	3.515	3.580
	∞	1.960	2.236	2.388	2.491	2.569	2.631	2.683	2.727	2.766	2.800	2.858	2.948	3.016	3.220	3.335	3.414	3.474
0.99	4	6.530	7.942	8.800	9.411	9.884	10.27	10.59	10.87	11.11	11.33	11.60	12.26	12.70	13.99	14.71	15.21	15.58
	5	5.044	5.072	6.530	6.940	7.253	7.509	7.725	7.911	8.074	8.219	8.462	8.857	9.154	10.05	10.55	10.89	11.15
	6	4.355	5.071	5.503	5.814	6.055	6.253	6.426	6.564	6.690	6.803	6.098	7.302	7.535	8.239	8.637	8.912	9.122
	7	3.963	4.562	4.922	5.181	5.382	5.546	5.685	5.806	5.912	6.006	6.162	6.425	6.621	7.217	7.558	7.791	7.979
	8	3.712	4.238	4.552	4.778	4.953	5.097	5.218	5.323	5.416	5.499	5.641	5.865	6.038	6.564	6.863	7.072	7.231
	9	3.537	4.014	4.297	4.560	4.657	4.787	4.896	4.990	5.074	5.148	5.277	5.479	5.634	6.110	6.362	6.572	6.717
	10	3.408	3.850	4.111	4.207	4.442	4.560	4.660	4.747	4.824	4.802	5.010	5.196	5.339	5.778	6.020	6.205	6.340
	15	3.074	3.426	3.631	3.776	3.888	3.979	4.056	4.123	4.182	4.234	4.325	4.468	4.573	4.916	5.112	5.249	5.354
	20	2.932	3.247	3.428	3.556	3.655	3.735	3.802	3.860	3.912	3.957	4.036	4.160	4.258	4.550	4.720	4.839	4.931
	25	2.862	3.146	3.317	3.435	3.526	3.600	3.663	3.716	3.763	3.806	3.878	3.992	4.079	4.348	4.503	4.612	4.696
	30	2.802	3.085	3.247	3.350	3.446	3.516	3.575	3.625	3.670	3.710	3.778	3.885	3.068	4.221	4.366	4.468	4.547
	60	2.684	2.939	3.062	3.182	3.258	3.319	3.371	3.415	3.454	3.488	3.547	3.639	3.710	3.925	4.048	4.134	4.200
	120	2.629	2.871	3.006	3.100	3.171	3.229	3.277	3.318	3.354	3.386	3.441	3.526	3.591	3.789	3.901	3.980	4.040
	∞	2.576	2.806	2.934	3.022	3.080	3.143	3.188	3.226	3.260	3.280	3.346	3.419	3.479	3.661	3.764	3.835	3.889

parameters are anticipated to have absolute limits and no excursions during the validation runs. Our discussion focuses on statistical interval-based acceptance criteria.

Interval-Based Criteria

As described, acceptance criteria can be set in a number of ways depending on the nature of the characteristic being tested (a continuous variable or attribute or a discrete variable). Purity data, obtained from chromatography, are generally assumed to follow a normal distribution. However, attribute data, such as bioburden counts (in colony forming units) for process buffers and pools, follow a different distribution, such as the Poisson distribution. Setting control limits based on such nonnormal distribution is done using a c chart, (which plots the number of defects or blemishes with control limits computed based on the Poisson distribution) (3).

Case Study

We use real data for our case study. The three statistical approaches we explore are applied to the yield data from one column used in a full-scale campaign. The data are shown in Table 1.

Table 3. Sampling distribution for student's t (t values)

Degrees of Freedom	10.00%[a] 20.00%[b]	5.00%[a] 10.00%[b]	2.50%[a] 5.00%[b]	1.00%[a] 2.00%[b]	0.50%[a] 1.00%[b]	0.25%[a] 0.50%[b]
1	3.078	6.314	12.706	25.452	63.656	127.321
2	1.886	2.920	4.303	6.205	9.925	14.089
3	1.638	2.353	3.182	4.177	5.841	7.453
4	1.533	2.132	2.776	3.495	4.604	5.598
5	1.476	2.015	2.571	3.163	4.032	4.773
6	1.440	1.943	2.447	2.969	3.707	4.317
7	1.415	1.895	2.365	2.841	3.499	4.029
8	1.397	1.860	2.306	2.752	3.355	3.833
9	1.383	1.833	2.262	2.685	3.250	3.690
10	1.372	1.812	2.228	2.634	3.169	3.581
11	1.363	1.796	2.201	2.593	3.106	3.497
12	1.356	1.782	2.179	2.560	3.055	3.428
13	1.350	1.771	2.160	2.533	3.012	3.372
14	1.345	1.761	2.145	2.510	2.977	3.326
15	1.341	1.753	2.131	2.490	2.947	3.286
16	1.337	1.746	2.120	2.473	2.921	3.252
17	1.333	1.740	2.110	2.458	2.898	3.222
18	1.330	1.734	2.101	2.445	2.878	3.197
19	1.328	1.729	2.093	2.433	2.861	3.174
20	1.325	1.725	2.086	2.423	2.845	3.153
21	1.323	1.721	2.080	2.414	2.831	3.135
22	1.321	1.717	2.074	2.405	2.819	3.119
23	1.319	1.714	2.069	2.398	2.807	3.104
24	1.318	1.711	2.064	2.391	2.797	3.091
25	1.316	1.708	2.060	2.385	2.787	3.078
26	1.315	1.706	2.056	2.379	2.779	3.067
27	1.314	1.703	2.052	2.373	2.771	3.057
28	1.313	1.701	2.048	2.368	2.763	3.047
29	1.311	1.699	2.045	2.364	2.756	3.038
30	1.310	1.697	2.042	2.360	2.750	3.030
31	1.309	1.696	2.040	2.356	2.744	3.022
32	1.309	1.694	2.037	2.352	2.738	3.015
40	1.303	1.684	2.021	2.329	2.704	2.971
50	1.299	1.676	2.009	2.311	2.678	2.937
60	1.296	1.671	2.000	2.299	2.660	2.915
120	1.289	1.658	1.980	2.270	2.617	2.860

[a]one-tailed probability (α value) for t to exceed table value
[b]two-tailed probability (α value) for t to exceed table value

Textbook method. Assuming normality, a confidence interval is defined by a standard deviation σ and a mean μ, and a multiplier k that represents the percentage point of a t- or z-distribution. So acceptance criteria are determined using the equality $\mu \pm k \times \sigma$ (referred to as Equation 1 hereafter). We propose three alternative statistical approaches to the textbook method of choosing $k = 3$ (that is, $\mu \pm 3 \times \sigma$). The methods are described in details in Hahn and Meeker (4).

Alternative 1: Prediction interval to contain all m future observations. The model calculates a $100(1-\alpha)\%$ prediction interval to contain m future runs based on n previous runs. Hahn and Meeker present tables (4) for different combinations of (m and n) at a given confidence level γ (Table 3). In our example, we estimate $\mu = 86.4$, an estimated $\sigma = 8.98394$. We want to compute a two-sided 95% prediction interval (α=5%) to contain five future observations (m=5) using 10 previous runs (n=10). Therefore, Equation 1 becomes $86.4 \pm 3.21 \times \sigma$ and $k = 3.321$ (see Table 2). A range such as that is reasonable in biotechnology processes. Tighter criteria might be appropriate, however, for characteristics directly related to final product quality (such as key purity

Table 4. Factors for calculating upper one-sided 95% and 99% statistical bounds for standard deviation of a normal distribution based on chi-square distribution (4, pp. 55, 296, Tables A.4a and A.4b)

Number of Observations (n)	Degrees of Freedom	95% Upper Bound	99% Upper Bound
4	3	2.92	5.11
5	4	2.37	3.67
6	5	2.09	3.00
7	6	1.92	2.62
8	7	1.80	2.38
9	8	1.71	2.20
10	9	1.65	2.08
12	11	1.55	1.90
15	14	1.46	1.73
20	19	1.37	1.58
25	24	1.32	1.49
30	29	1.28	1.43
40	39	1.23	1.35
60	59	1.18	1.27
∞	∞	1.00	1.00

indicators) to prevent conflicts with CFR 211.110b (1).

Alternative 2: Confidence interval for the standard deviation (s) combined with the confidence interval for the mean of a normal distribution. We can use chi-square (χ^2) distribution to determine the one-sided 100(1−α)% upper limit on σ (4).

Table 5. Factors for calculating normal distribution two-sided 100(1−α)% tolerance intervals to simultaneously control each tail of the distribution to 100*p*% or less; all values are for 95% confidence (α=5%) so that *p* portion of the population will be simultaneously in each tail of the distribution (4, Pg. 59, 311, Table A.11b)

Number of Observations (n)	p=0.05 (1−α)=0.95	p=0.025 (1−α)=0.95
2	35.225	40.251
3	9.408	10.785
4	6.074	6.980
5	4.847	5.582
6	4.209	4.855
7	3.815	4.407
8	3.546	4.100
9	3.348	3.876
10	3.197	3.704
11	3.076	3.568
12	2.978	3.456
13	2.895	3.363
14	2.825	3.284
15	2.765	3.216
16	2.713	3.157
17	2.666	3.104
18	2.625	3.058
19	2.588	3.016
20	2.555	2.978
21	2.524	2.944
22	2.497	2.913
23	2.471	2.884
24	2.448	2.858
25	2.426	2.833
26	2.406	2.811
27	2.387	2.790
28	2.370	2.770
29	2.353	2.752
30	2.338	2.734
35	2.273	2.661
40	2.223	2.605
50	2.149	2.522
60	2.097	2.464
120	1.949	2.298
240	1.853	2.191
480	1.788	2.119
∞	1.645	1.960

The result of such an operation is a larger σ, denoted by σ^*. The formula (Equation 1) becomes $\mu \pm k \times \sigma^*$, where k is derived from a t-distribution table with $n-1$ degrees-of-freedom. Table 3 shows the sampling distribution for the t-statistic, and

USING ranges slightly wider than the conventional "mean ± 3 standard deviations" will demonstrate that you recognize the importance of sample size in setting acceptance criteria.

Table 4 gives factors based on the χ^2 distribution for calculating the statistical bounds for the standard deviation of a normal distribution. Another way of using this formula is to define a new k using χ^2 distribution *and* t-distribution simultaneously denoted by k^*, so that

$$k^* = t\left(\frac{1-\alpha}{2; n-1}\right)\left(\frac{n-1}{\chi^2(\alpha; n-1)}\right)^{\frac{1}{2}}$$

Our model for Equation 1 is then $\mu \pm k^* \times \sigma$. As the sample size increases, the factor of the confidence boundary (or confidence limit) decreases. In our example, for $n=10$, a 95% one-sided upper limit on σ will use the multiplication factor k^*:

$$k^* = \left(\frac{n-1}{\chi^2_{(\alpha; n-1)}}\right)^{\frac{1}{2}} = 1.65$$

From the t-distribution: $t\left(\frac{1-\alpha}{2}; n-1\right) = 2.262$ for two-sided 95% confidence. The formula in Equation 1 becomes:
$86.4 \pm 2.262 \times 1.65 \times \sigma$
or $86.4 \pm 3.732 \times \sigma$.

Alternative 3: Tolerance Intervals to contain a proportion of a population. This alternative is interval-based to contain on the average a certain proportion *p* of the population. Such an interval becomes the *tolerance interval* (4). Applied to our example, the task is to estimate tolerance limits that will on the average contain, for example, 95% of the population at 95% confidence. Hahn and Meeker offer appropriate tables (4, Tables A:10b and A:11b, pg. 59) for combinations of *p*, *n*, and confidence level γ. Table 10b is used when the investigator wants to control the center of the distribution, and Table 11b is used when the interest is to control both tails of the distribution. Because we wish to control the proportion of the population in the tails of the distribution, we have reproduced an appropriate segment of Table 11b as our Table 5. For γ=0.95, *p*=0.95, and *n*=10 Table 5 gives a k=3.197. Using those results on the same data gives μ=86.4 and σ=8.98394, which means the tolerance limits are obtained by $86.4 \pm 3.197 \times \sigma$.

Recognizing Sample Size

For reasons of simplicity and the fact that alternative 2 can be seen as an extension of the textbook method, we recommend alternative 2 as the best approach. However, alternatives 1 and 3 can be appropriate for certain situations. Both the textbook method and alternative 2 use a confidence interval based on Equation 1, $\mu \pm k \times \sigma$.

The textbook method uses k=3 (that is 3σ) and makes no allowance for the uncertainty in the estimate of σ or of a μ based on small sample sizes. Alternatives 1, 2, and 3 recognize the importance of sample size in setting acceptance criteria.

References

(1) *Code of Federal Regulations, Food and Drugs*, "Current Good Manufacturing Practice in Manufacturing, Processing, Packing, or Holding of Drugs," Title 21, Part 211 (U.S. Government Printing Office, Washington, DC), revised 1 April 2001.

(2) D.A. Greer and E.J. Halteman, "The Swiss Army Knife of Control Charts," paper presented at the 1993 Joint Meetings of the American Statistical Association, San Francisco, 8–12 August 1993.

(3) D.J. Wheeler and D.S. Chambers, *Understanding Statistical Process Control*, 2nd Ed. (SPC Press, Knoxville, TN, 1992), p. 271.

(4) G.J. Hahn and W.Q. Meeker, "Prediction Interval to Contain All of M Future Observations," *Statistical Intervals: A Guide for Practitioners* (John Wiley & Sons, NY, 1991), pp. 55, 59, 62. **BP**