PET ENGINEERING COLLEGE

An ISO 9001:2015 Certified Institution

Accredited by NAAC, Approved by AICTE, Recognized by Government of Tamil Nadu and Affiliated to Anna University

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

UNIT – II

YIELD, MEASUREMENT ACCURACY, AND TEST TIME

CLASS : S6 ECE

SUBJECT CODE : CEC342

SUBJECT NAME : MIXED SIGNAL IC DESIGN TESTING

REGULATION : 2021

CHAPTER - II

YIELD, MEASUREMENT, ACCURACY & TEST TIME

Testing is an important and essential phase in the manufacturing process of integrated circuits. In fact, the only way that manufacturers can deliver high-quality ICs in a reasonable time is through clever testing procedures. The IC manufacturing process involves three major steps: fabrication, testing, and packaging. Today, manufacturing costs associated with mixed-signal ICs is being dominated by the test phase (i.e., separating bad dies from good ones), although packaging costs are becoming quite significant in some large ICs. In order to create clever test procedures, one needs to have a clear understanding of the tradeoffs involved. In particular, the test engineer needs to understand the needs of their business (making ICs for toys or the automotive industry), the cost of test, and the quality of the product produced. It is the intent of this chapter to outline these tradeoffs, beginning with a discussion on manufacturing yield, followed by a discussion on measurement accuracy, and then moving to a discussion of test time. As the needs of a business is highly variable, we will make comments throughout this chapter where appropriate.

5.1 YIELD

The primary goal of a semiconductor manufacturer is to produce large quantities of ICs for sale to various electronic markets-that is, cell phones, ipods, HDTVs, and so on. Semiconductor factories are highly automated, capable of producing millions of ICs over a 24-hour period, every day of the week. For the most part, these ICs are quite similar in behavior, although some will be quite different from one another. A well-defined means to observe the behavior of a set of large elements, such as ICs, is to categorize their individual behavior in the form of a histogram, as shown in Figure 5.1. Here we illustrate a histogram of the offset voltage associated with a lot of devices. We see that 15% of devices produced in this lot had an offset voltage between -0.129 V and -0.128 V. We can conjecture that the probability of another lot producing devices with an offset voltage in this same range is 15%. Of course, how confident we are with our conjecture is the basis of all things statistical; we need to capture more data to support our claim. This we will address shortly; for now, let us consider the "goodness" of what we produced.

In general, the component data sheet defines the "goodness" of an analog or mixed-signal device. As a data sheet forms the basis of any contract between a supplier and a buyer, we avoid any subjective argument of why one measure is better or worse than another; it is simply a matter of data sheet definition. Generally, the goodness of an analog and mixed-signal device is defined by a range of acceptability, bounded by a lower specification limit (LSL) and an upper specification limit (USL), as further illustrated in Figure 5.1. These limits would be found on the device data sheet. Any device whose behavior falls outside this range would be considered as a bad device. This particular example considers a device with a two-sided limit. Similarly, the same argument applies to a one-sided limit; just a different diagram is used.

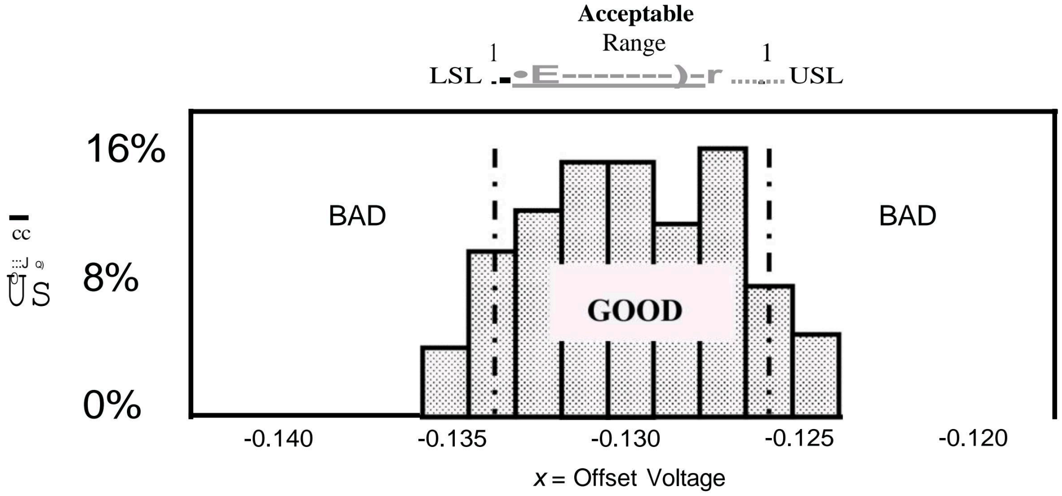
Testing is the process of separating good devices from the bad ones. The yield of a given lot of material is defined as the ratio of the total good devices divided by the total devices tested:

$$\text{yield} = \frac{\text{total good devices}}{\text{total devices tested}} \times 100\% \tag{5.1}$$

If 10,000 parts are tested and only 7000 devices pass all tests, then the yield on that lot of 10,000 devices is 70%. Because testing is not a perfect process, mistakes are made, largely on account of the measurement limitations of the tester, noise picked up at the test interface, and noise produced by the DUT itself. The most critical error that can be made is one where a bad device is declared good, because this has a direct impact on the operations of a buyer. This error is known as an escape. As a general rule, the impact that an escape has on a manufacturing process goes up exponentially as it moves from one assembly level to another. Hence, the cost of an escape can be many orders of magnitude greater than the cost of a single part. Manufacturers make use of test metrics to gauge the goodness of the component screening process. One measure is the defect level (DL) and it is defined as

$$DL = \frac{\text{total bad devices declared good}}{\text{total devices declared good}} \times 100\% \tag{5.2}$$

Figure 5.1. Histogram showing specification limits and regions of acceptance and rejection.



or, when written in terms of escapes, we write

$$DL = \frac{\text{total escapes}}{\text{total devices declared good}} \times 100\% \tag{5.3}$$

Often DL is expressed in the number of defects-per-million or what is more commonly stated as parts per million (ppm).

It is important to note that a measure of defect level is a theoretical concept based on a probability argument and one that has no empirical basis, because if we knew which devices were escapes, then we would be able to identify them as bad and remove them from the set of good devices. Nonetheless, companies do estimate their defect levels from various measures based on their field returns, or by analyzing the test data during the test process using a secondary screening procedure.

Exercises

5.1. If 15,000 devices are tested with a yield of 63%, how many devices passed the test? ANS. 9450 devices.

5.2. A new product was launched with 100,000 sales over a one-year time span. During this time, seven devices were returned to the manufacturer even though an extensive test screening procedure was in place. What is the defect level associated with this testing procedure in parts per million? ANS. 70 ppm.

5.2 MEASUREMENT TERMINOLOGY

The cause of escapes in an analog or mixed-signal environment is largely one that is related to the measurement process itself. A value presented to the instrument by the DUT will introduce errors, largely on account of the electronic circuits that make up the instrument. These errors manifest themselves in various forms. Below we shall outline these errors first in a qualitative way, and then we will move onto a quantitative description in the next section on how these errors interrelate and limit the measurement process.

5.2.1 Accuracy and Precision

In conversational English, the terms *accuracy* and *precision* are virtually identical in meaning. *Roget's Thesaurus*[1] lists these words as synonyms, and *Webster's Dictionary*[2] gives almost identical definitions for them. However, these terms are defined very differently in engineering textbooks.[3-5] Combining the definitions from these and other sources gives us an idea of the accepted technical meaning of the words:

Accuracy: The difference between the average of measurements and a standard sample for which the "true" value is known. The degree of conformance of a test instrument to absolute standards, usually expressed as a percentage of reading or a percentage of measurement range (full scale).

Precision: The variation of a measurement system obtained by repeating measurements on the same sample back-to-back using the same measurement conditions.

According to these definitions, precision refers only to the repeatability of a series of measurements. It does not refer to consistent errors in the measurements. A series of measurements can be incorrect by 2 V, but as long as they are consistently wrong by the same amount, then the measurements are considered to be precise.

This definition of precision is somewhat counterintuitive to most people, since the words *precision* and *accuracy* are so often used synonymously. Few of us would be impressed by a "precision" voltmeter exhibiting a consistent 2-V error! Fortunately, the word *repeatability* is far more commonly used in the test-engineering field than the word *precision.* This textbook will use the term *accuracy* to refer to the overall closeness of an averaged measurement to the true value and *repeatability* to refer to the consistency with which that measurement can be made. The word *precision* will be avoided.

Unfortunately, the definition of accuracy is also somewhat ambiguous. Many sources of error can affect the accuracy of a given measurement. The accuracy of a measurement should probably refer to all possible sources of error. However, the accuracy of an instrument (as distinguished from the accuracy of a measurement) is often specified in the absence of repeatability fluctuations and instrument resolution limitations. Rather than trying to decide which of the various error sources are included in the definition of accuracy, it is probably more useful to discuss some of the common error components that contribute to measurement inaccuracy. It is incumbent upon the test engineer to make sure all components of error have been accounted for in a given specification of accuracy.

5.2.2 Systematic or Bias Errors

Systematic or bias errors are those that show up consistently from measurement to measurement. For example, assume that an amplifier's output exhibits an offset of 100 mV from the ideal value of O V. Using a digital voltmeter (DVM), we could take multiple readings of the offset over time and record each measurement. A typical measurement series might look like this:

101 mV, 103 mV, 102 mV, 101 mV, 102 mV, 103 mV, 103 mV, 101 mV, 102 mV ...

This measurement series shows an average error of about 2 mV from the true value of 100 mV. Errors like this are caused by consistent errors in the measurement instruments. The errors can result from a combination of many things, including DC offsets, gain errors, and nonideal linearity in the DVM's measurement circuits. Systematic errors can often be reduced through a process called *calibration.* Various types of calibration will be discussed in more detail in Section 5.4.

5.2.3 Random Errors

In the preceding example, notice that the measurements are not repeatable. The DVM gives readings from 101 to 103 mV. Such variations do not surprise most engineers because DVMs are relatively inexpensive. On the other hand, when a two million dollar piece of ATE equipment cannot produce the same answer twice in a row, eyebrows may be raised.

Inexperienced test engineers are sometimes surprised to learn that an expensive tester cannot give perfectly repeatable answers. They may be inclined to believe that the tester software is defective when it fails to produce the same result every time the program is executed. However, experienced test engineers recognize that a certain amount of random error is to be expected in analog and mixed-signal measurements.

Random errors are usually caused by thermal noise or other noise sources in either the DUT or the tester hardware. One of the biggest challenges in mixed-signal testing is determining whether the random errors are caused by bad DIB design, by bad DUT design, or by the tester itself. If the source of error is found and cannot be corrected by a design change, then averaging or filtering of measurements may be required. Averaging and filtering are discussed in more detail in Section 5.6.

5.2.4 Resolution (Quantization Error)

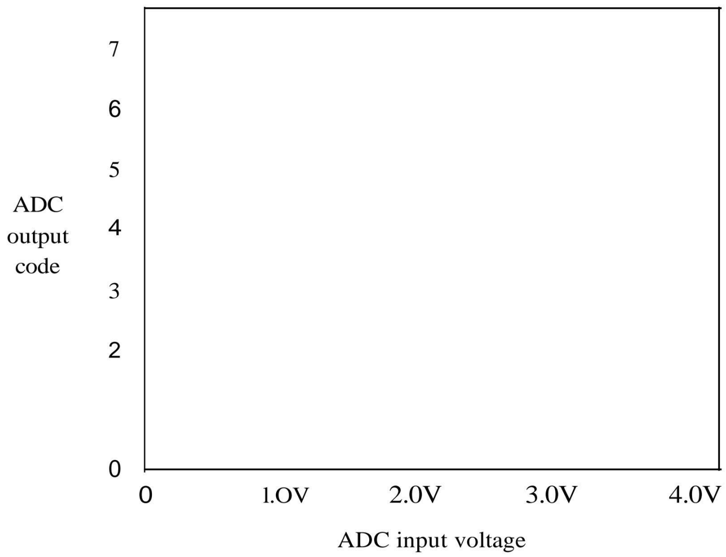
In the 100-mV measurement list, notice that the measurements are always rounded off to the nearest millivolt. The measurement may have been rounded off by the person taking the measurements, or perhaps the DVM was only capable of displaying three digits. ATE measurement instruments have similar limitations in measurement resolution. Limited resolution results from the fact that continuous analog signals must first be converted into a digital format before the ATE computer can evaluate the test results. The tester converts analog signals into digital form using analog-to-digital converters (ADCs).

ADCs by nature exhibit a feature called *quantization error*. Quantization error is a result of the conversion from an infinitely variable input voltage (or current) to a finite set of possible digital output results from the ADC. Figure 5.2 shows the relationship between input voltages and output codes for an ideal 3-bit ADC. Notice that an input voltage of 1.2 V results in the same ADC output code as an input voltage of 1.3 V. In fact, any voltage from 1.0 to 1.5 V will produce an output code of 2.

If this ADC were part of a crude DC voltmeter, the meter would produce an output reading of 1.25 V any time the input voltage falls between 1.0 and 1.5 V. This inherent error in ADCs and measurement instruments is caused by quantization error. The resolution of a DC meter is often limited by the quantization error of its ADC circuits.

If a meter has 12 bits of resolution, it means that it can resolve a voltage to one part in $2^{12} - 1$ (one part in 4095). If the meter's full-scale range is set to ±2 V, then a resolution of approximately 1 mV can be achieved (4 V/4095 levels). This does not automatically mean that the meter is accurate to 1 mV, it simply means the meter cannot resolve variations in input voltage smaller than 1 mV. An instrument's resolution can far exceed its accuracy. For example, a 23-bit voltmeter might be able to produce a measurement with a 1-μV resolution, but it may have a systematic error of 2 mV.

Figure 5.2. Output codes versus input voltages for an ideal 3-bit ADC.



5.2.5 Repeatability

Nonrepeatable answers are a fact oflife for mixed-signal test engineers. A large portion of the time required to debug a mixed-signal test program can be spent tracking down the various sources of poor repeatability. Since all electrical circuits generate a certain amount of random noise, measurements such as those in the 100-mV offset example are fairly common. In fact, if a test engineer gets the same answer 10 times in a row, it is time to start looking for a problem. Most likely, the tester instrument's full-scale voltage range has been set too high, resulting in a measurement resolution problem. For example, if we configured a meter to a range having a 10-mV resolution, then our measurements from the prior example would be very repeatable (100 mV, 100 mV, 100 mV, 100 mV, etc.). A novice test engineer might think that this is a terrific result, but the meter is just rounding off the answer to the nearest 10-mV increment due to an input ranging problem. Unfortunately, a voltage of 104 mV would also have resulted in this same series of perfectly repeatable, perfectly incorrect measurement results. Repeatability is desirable, but it does not in itself guarantee accuracy.

5.2.6 Stability

A measurement instrument's performance may drift with time, temperature, and humidity. The degree to which a series of supposedly identical measurements remains constant over time, temperature, humidity, and all other time-varying factors is referred to as *stability.* Stability is an essential requirement for accurate instrumentation.

Shifts in the electrical performance of measurement circuits can lead to errors in the tested results. Most shifts in performance are caused by temperature variations. Testers are usually equipped with temperature sensors that can automatically determine when a temperature shift has occurred. The tester must be recalibrated anytime the ambient temperature has shifted by a few degrees. The calibration process brings the tester instruments back into alignment with known electrical standards so that measurement accuracy can be maintained at all times.

After the tester is powered up, the tester's circuits must be allowed to stabilize to a constant temperature before calibrations can occur. Otherwise, the measurements will drift over time as the tester heats up. When the tester chassis is opened for maintenance or when the test head is opened up or powered down for an extended period, the temperature of the measurement electronics will typically drop. Calibrations then have to be rerun once the tester recovers to a stable temperature.

Shifts in performance can also be caused by aging electrical components. These changes are typically much slower than shifts due to temperature. The same calibration processes used to account for temperature shifts can easily accommodate shifts of components caused by aging. Shifts caused by humidity are less common, but can also be compensated for by periodic calibrations.

5.2.7 Correlation

Correlation is another activity that consumes a great deal of mixed-signal test program debug time. Correlation is the ability to get the same answer using different pieces of hardware or software. It can be extremely frustrating to try to get the same answer on two different pieces of equipment using two different test programs. It can be even more frustrating when two supposedly identical pieces of test equipment running the same program give two different answers.

Of course correlation is seldom perfect, but how good is good enough? In general, it is a good idea to make sure that the correlation errors are less than one-tenth of the full range between the minimum test limit and the maximum test limit. However, this is just a rule of thumb. The exact requirements will differ from one test to the next. Whatever correlation errors exist, they must be considered part of the measurement uncertainty, along with nonrepeatability and systematic errors.

The test engineer must consider several categories of correlation. Test results from a mixed-signal test program cannot be fully trusted until the various types of correlation have been verified. The more common types of correlation include tester-to-bench, tester-to-tester, program-to-program, DIB-to-DIB, and day-to-day correlation.

Tester-to-Bench Correlation

Often, a customer will construct a test fixture using bench instruments to evaluate the quality of the device under test. Bench equipment such as oscilloscopes and spectrum analyzers can help validate the accuracy of the ATE tester's measurements. Bench correlation is a good idea, since ATE testers and test programs often produce incorrect results in the early stages of debug. In addition, IC design engineers often build their own evaluation test setups to allow quick debug of device problems. Each of these test setups must correlate to the answers given by the ATE tester. Often the tester is correct and the bench is not. Other times, test program problems are uncovered when the ATE results do not agree with a bench setup. The test engineer will often need to help debug the bench setup to get to the bottom of correlation errors between the tester and the bench.

Tester-to-Tester Correlation

Sometimes a test program will work on one tester, but not on another presumably identical tester. The differences between testers may be catastrophically different, or they may be very subtle. The test engineer should compare all the test results on one tester to the test results obtained using other testers. Only after all the testers agree on all tests is the test program and test hardware debugged and ready for production.

Similar correlation problems arise when an existing test program is ported from one tester type to another. Often, the testers are neither software compatible nor hardware compatible with one another. In fact, the two testers may not even be manufactured by the same ATE vendor. A myriad of correlation problems can arise because of the vast differences in DIB layout and tester software between different tester types. To some extent, the architecture of each tester will determine the best test methodology for a particular measurement. A given test may have to be executed in a very different manner on one tester versus another. Any difference in the way a measurement is taken can affect the results. For this reason, correlation between two different test approaches can be very difficult to achieve. Conversion of a test program from one type of tester to another can be one of the most daunting tasks a mixed-signal test engineer faces.

Program-to-Program Correlation

When a test program is streamlined to reduce test time, the faster program must be correlated to the original program to make sure no significant shifts in measurement results have occurred. Often, the test reduction techniques cause measurement errors because of reduced DUT settling time and other timing-related issues. These correlation errors must be resolved before the faster program can be released into production.

D/8-to-DIB Correlation

No two DIBs are identical, and sometimes the differences cause correlation errors. The test engineer should always check to make sure that the answers obtained on multiple DIB boards agree. DIB correlation errors can often be corrected by *focused calibration* software written by the test engineer (this will be discussed further in Section 5.4).

Exercises

5.3. A 5-mV signal is measured with a meter 10 times, resulting in the following sequence of readings: 5 mV, 6 mV, 9 mV, 8 mV, 4 mV, 7 mV, 5 mV, 7 mV, 8 mV, 11 mV. What is the average measured value? What is the systematic error? ANS. 7 mV, 2 mV.

5.4. A meter is rated at 8-bits and has a full-scale range of ±5 V. What is the measurement uncertainty of this meter, assuming only quantization errors from an ideal meter ADC? ANS. ±19.6 mV.

5.5. A signal is to be measured with a maximum uncertainty of ±0.5 µV. How many bits of resolution are required by an ideal meter having a ±1 V full-scale range? ANS. 21 bits.

Day-to-Day Correlation

Correlation of the same DIB and tester over a period of time is also important. If the tester and DIB have been properly calibrated, there should be no drift in the answers from one day to the next. Subtle errors in software and hardware often remain hidden until day-to-day correlation is performed. The usual solution to this type of correlation problem is to improve the focused calibration process.

5.2.8 Reproducibility

The term *reproducibility* is often used interchangeably with *repeatability,* but this is not a correct usage of the term. The difference between reproducibility and repeatability relates to the effects of correlation and stability on a series of supposedly identical measurements. Repeatability is most often used to describe the ability of a single tester and DIB board to get the same answer multiple times as the test program is repetitively executed.

Reproducibility, by contrast, is the ability to achieve the same measurement result on a given DUT using any combination of equipment and personnel at any given time. It is defined as the statistical deviation of a series of supposedly identical measurements taken over a period of time. These measurements are taken using various combinations of test conditions that ideally should not change the measurement result. For example, the choice of equipment operator, tester, DIB board, and so on, should not affect any measurement result.

Consider the case in which a measurement is highly repeatable, but not reproducible. In such a case, the test program may consistently pass a particular DUT on a given day and yet consistently fail the same DUT on another day or on another tester. Clearly, measurements must be both repeatable and reproducible to be production-worthy.

5.3 A MATHEMATICAL LOOK AT REPEATABILITY, BIAS, AND ACCURACY

To gain a better understanding of the meaning of accuracy and its impact on a measurement,[6] consider the circuit diagram of a voltage-reference-voltmeter arrangement shown in Figure 5.3. Here we model the reference level with a DC voltage source with value V_{REF}. The voltmeter is modeled with an ideal meter with reading $V_{MEASURED}$ and two voltage sources in series with the reference. One voltage source represents the offset introduced by the voltmeter (V_{OFF}), and the other represents the noise generated by the voltmeter (V_{noise}). By KVL, we can write the voltmeter value as

$$V_{MEASURED} = V_{REF} + V_{OFF} + V_{noise} \quad (5.4)$$

If we repeat a sequence of measurements involving the same reference, we would obtain a set of values that would in general be all different on account of the noise that is present. To eliminate the effects of this noise, one could instead take the average value of a large number of samples as the measurement. For instance, if we take the expected or average value of each side of Eq. (5.4), we write

$$E[V_{MEASURED}] = E[V_{REF} + V_{OFF} + V_{noise}] \quad (5.5)$$

Recognizing that the expectation operation distributes across addition, we can write

$$E[V_{MEASURED}] = E[V_{REF}] + E[V_{OFF}] + E[V_{noise}] \quad (5.6)$$

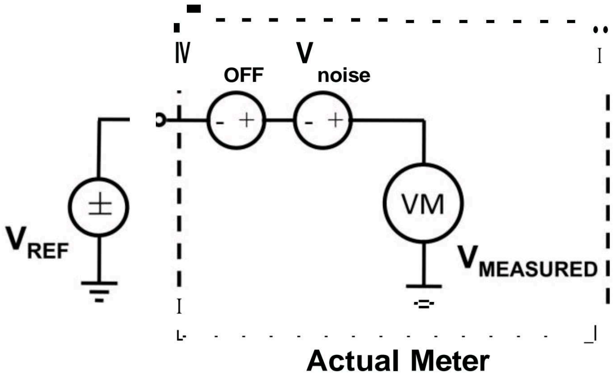
Assuming that the noise process is normal with zero mean, together with the fact that V_{REF} and V_{OFF} are constants, we find that the expected measured value becomes

$$E[V_{MEASURED}] = V_{REF} + V_{OFF} \quad (5.7)$$

As long as the sample set is large, then averaging will eliminate the effects of noise. However, if the sample size is small, a situation that we often find in practice, then our measured value will vary from one sample set to another. See, for example, the illustration in Figure 5.4a involving two sets of samples. Here we see the mean values μ_{M1} and μ_{M2} are different. If we increase the total number of samples collected to, say, N, we would find that the mean values of the two distributions approach one another in a statistical sense. In fact, the mean of the means will converge to $V_{REF} + V_{OFF}$ with a standard deviation of $\sigma_M / \sqrt{N}$ as illustrated by the dashed line distribution shown in Figure 5.4b. We should also note that the distribution of means is indeed Gaussian as required by the central limit theorem.

Metrology (the science of measurement) is interested in quantifying the level of uncertainty present in a measurement. Three terms from metrology are used in test to describe this uncertainty: repeatability, accuracy, and bias.

Figure 5.3. Modeling a measurement made by a voltmeter with offset and noise component.



Assume that N samples are taken during some measurement process and that these samples are assigned to vector x. The mean value of the measurement is quantified as

$$\mu_M = \frac{1}{N}\sum_{k=1}^{N} x[k] \tag{5.8}$$

The repeatability of a measurement refers to the standard deviation associated with the measurement set, that is,

$$\sigma_M = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(x[k]-\mu_M\right)^2} \tag{5.9}$$

For the example shown in Figure 5.4b, repeatability refers to the spread of the measurement samples about the sample mean value - The larger the spread, the less repeatable the measurement will be. We can now define repeatability as the variation (quantified by σ_M) of a measurement system obtained by repeating measurements on the same sample back-to-back using the same measurement conditions.

Bias error or systematic error is the difference between the reference value and the average of a large number of measured values. Bias error can be mathematically described as

$$\beta = v_{REF} - E[y_{MEASURED}] \tag{5.10}$$

where $E[y_{MEASURED}]$ is derived through a separate measurement process involving a (very) large number of samples, that is,

$$E[y_{MEASURED}] = \frac{1}{N}\sum_{k=1}^{N} x[k]\Bigg|_{N\ LARGE} \tag{5.11}$$

This step is usually conducted during the characterization phase of the product rather than during a production run to save time. In essence, $E[y_{MEASURED}]$ converges to $v_{REF} + v_{OFF}$ (the noiseless value) and equals the negative of the instrument offset, that is,

$$\beta = -v_{OFF} \tag{5.12}$$

Finally, we come to the term *accuracy*. Since test time is of critical importance during a production test, the role of the test engineer is to make a measurement with just the right amount of uncertainty-no more, no less. This suggests selecting the test conditions so that the accuracy of the measurement is just right.

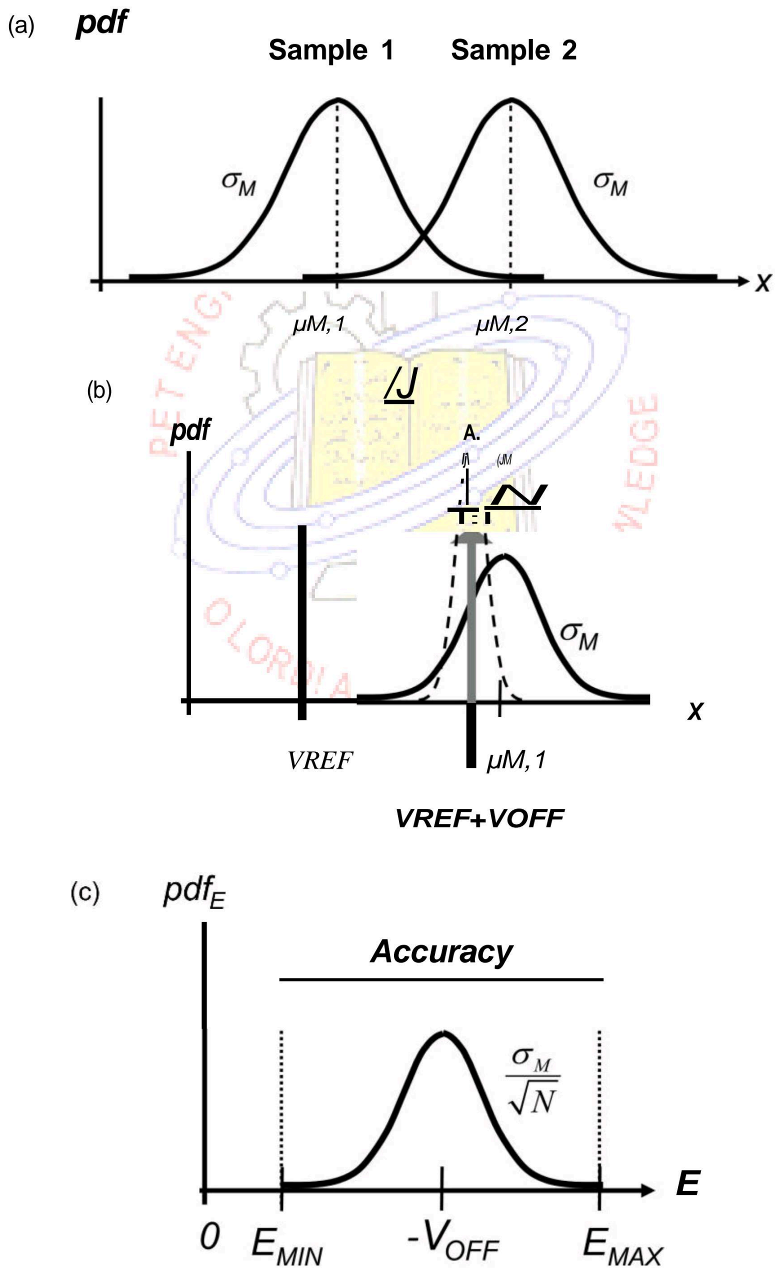
Like bias error, accuracy is defined in much the same way, that being the difference between the known reference value and the expected value of the measurement process. However, accuracy accounts for the error that is introduced due to the repeatability of the measurement that is caused by the small sample set. Let us define the difference between the reference level v_{REF} and an estimate of the mean value, given by $\overline{v}_{MEASURED} = \frac{1}{N}\sum_{k=1}^{N} x[k]\Big|_{N\ SMALL}$, as the measurement error:

$$E = V_{REF} - V_{MEASURED} \tag{5.13}$$

Absolute accuracy of a measurement is then defined in terms of a range of possible errors as

$$E_{MIN} \le \text{accuracy} \le E_{MAX} \tag{5.14}$$

Figure 5.4. (a) Small sets of different measurements will have different mean values. (b) The mean value of a large sample set will converge to VREF with an offset V_{OFF}. (c) Distribution of measurement errors.



where

$$E_{MIN} = \min\{V_{REF} - \min[V_{MEASURED}],\ V_{REF} - \max[V_{MEASURED}]\}$$

$$E_{MAX} = \max\{V_{REF} - \min[V_{MEASURED}],\ V_{REF} - \max[V_{MEASURED}]\}$$

It is common practice to refer to the range bounded by E_{MAX} and E_{MIN} as the uncertainty range of the measurement process, or simply just accuracy, as shown in Figure 5.4c. Clearly, some measure of the distribution of measurement errors must be made to quantify accuracy. We have more to say about this in a moment.

As a reference check, if the measurement process is noiseless, then we have

$$\max[V_{MEASURED}] = \min[V_{MEASURED}] = V_{REF} + V_{OFF} \tag{5.15}$$

and the accuracy of the measurement process would simply be equal to the offset term, that is,

$$\text{accuracy} = -V_{OFF} \tag{5.16}$$

At this point in the discussion of accuracy, it is important to recognize that measurement offset plays an important role in the role of accuracy in a measurement. We have a lot more to say about this in a moment.

Absolute accuracy is also expressed in terms of the center value and a plus-minus difference measure defined as

$$\text{accuracy} = \frac{E_{MAX} + E_{MIN}}{2} \pm \frac{|E_{MAX} - E_{MIN}|}{2} \tag{5.17}$$

It is important that the reader be aware of the meaning of the notation used in Eq. (5.17) and how it maps to the error bound given by Eq. (5.14).

Sometimes accuracy is expressed in relative terms, such as a percentage of reference value, that is,

$$\text{accuracy}_{\%} = \left[\frac{1}{2}\frac{E_{MAX} + E_{MIN}}{V_{REF}} \pm \frac{1}{2}\left|\frac{E_{MAX} - E_{MIN}}{V_{REF}}\right|\right] \times 100\% \tag{5.18}$$

or a percentage of the full-scale measurement range (denote V_{FS}) as

$$\text{accuracy}_{\%FS} = \left[\frac{1}{2}\frac{E_{MAX} + E_{MIN}}{V_{FS}} \pm \frac{1}{2}\left|\frac{E_{MAX} - E_{MIN}}{V_{FS}}\right|\right] \times 100\% \tag{5.19}$$

EXAMPLE 5.1

A measurement of a 1-V reference level is made 100 times, where the minimum reading is 0.95 V and the maximum reading is 1.14 V. What is the absolute accuracy of these measurements? What is the relative accuracy with respect to a full-scale value of 5 V?

Solution:

According to Eq. [5.14]. the smallest and largest errors are

$$E_{MIN} = \min\{1-0.95, 1-1.14\} = -0.14$$
$$E_{MAX} = \max\{1-0.95, 1-1.14\} = 0.05$$

leading to an absolute accuracy bound of

$$-0.14 \le \text{accuracy} \le 0.05$$

or expressed with respect to the center as

$$\text{accuracy} = \frac{-0.14+0.05}{2} \pm \frac{|-0.14-0.05|}{2} = -0.045 \pm 0.095\text{V}$$

In terms of a relative accuracy referenced to the 5-V full-scale level, we obtain

$$\text{accuracy}_{\%FS} = \frac{-0.045 \pm 0.095\text{V}}{5\text{V}} \times 100\% = -0.9 \pm 1.9\%$$

It is interesting to observe the statistics of this estimator, $\hat{V}_{MEASURED}$, when noise is present. Regardless of the nature of the noise, according to the central limit theorem, the estimator $\hat{V}_{MEASURED}$ will follow Gaussian statistics. This is because the estimator is a sum of several random variables as described in the previously chapter. The implications of this is that if one were to repeat a measurement a number of times and create a histogram of resulting estimator values, one would find that it has a Gaussian shape. Moreover, it would have a standard deviation of approximately $\sigma_M/\sqrt{N}$. As a first guess, we could place the center of the Gaussian distribution at the value of the estimator $\hat{V}_{MEASURED} = \frac{1}{N}\sum_{k=1}^{N} x[k]$. Hence, the pdf of the estimator values would appear as

$$g(v) = \frac{\sqrt{N}}{\sigma_M\sqrt{2\pi}} e^{\frac{-\left(v-\hat{V}_{MEASURED}\right)^2}{2\left(\sigma_M/\sqrt{N}\right)^2}} \tag{5.20}$$

Consequently, we can claim with a 99.7% probability that the true mean of the measurement will lie between $\overline{V}_{MEASURED} - 3\frac{\sigma_M}{\sqrt{N}}$ and $\overline{V}_{MEASURED} + 3\frac{\sigma_M}{\sqrt{N}}$. One can introduce an a term to the previous range term and generalize the result to a set of probabilities-for example,

$$P\left(\overline{V}_{MEASURED} - a\frac{\sigma_M}{\sqrt{N}} \le E\left[V_{MEASURED}\right] \le \overline{V}_{MEASURED} + a\frac{\sigma_M}{\sqrt{N}}\right) = \begin{cases} 0.667, & a = 1 \\ 0.950, & a = 2 \\ 0.997, & a = 3 \end{cases} \tag{5.21}$$

One can refer to the a term as a confidence parameter, because the larger its value, the greater our confidence (probability) that the noiseless measured value lies within the range defined by

$$\overline{V}_{MEASURED} - a\frac{\sigma_M}{\sqrt{N}} \le E\left[V_{MEASURED}\right] \le \overline{V}_{MEASURED} + a\frac{\sigma_M}{\sqrt{N}} \tag{5.22}$$

In the statistical literature this range is known as the confidence interval (CI). The extremes of measurement estimator can then be identified as

$$\begin{aligned} \max\left[\overline{V}_{MEASURED}\right] &= \overline{V}_{MEASURED} + a\frac{\sigma_M}{\sqrt{N}} \\ \min\left[\overline{V}_{MEASURED}\right] &= \overline{V}_{MEASURED} - a\frac{\sigma_M}{\sqrt{N}} \end{aligned} \tag{5.23}$$

where $\overline{V}_{MEASURED}$ is any one estimate of the mean of the measurement and σ_M is the standard deviation of the measurement process usually identified during a characterization phase. Substituting Eq. (5.23) into Eq. (5.17), together with definitions given in Eq. (5.14), we write the accuracy expression as

$$\text{accuracy} = V_{REF} - \overline{V}_{MEASURED} \pm a\frac{\sigma_M}{\sqrt{N}} \tag{5.24}$$

As a first-order approximation, let us assume

$$V_{REF} - \overline{V}_{MEASURED} \approx \beta \tag{5.25}$$

Then Eq. (5.24) becomes

$$\text{accuracy} = \beta \pm a\frac{\sigma_M}{\sqrt{N}} \tag{5.26}$$

This is the **fundamental** equation for **measurement accuracy.** It illustrates the dependency of accuracy on the bias error, repeatability, and the number of samples. It also suggest several ways in which to improve measurement accuracy:

1. Remove the bias error β by calibrating to a known reference value.
2. Decrease the intrinsic amount of noise σ_M associated with a measurement by purchasing more expensive instruments with a lower noise floor or by improving the device interface board **(DIB)** design and test interface.
3. Increase the size of the sample set N to reduce the influence of measurement repeatability; increase the time of test; or alter the algorithm that is used to extract the mean value.

The next few sections will address these three points in greater detail.

EXAMPLE 5.2

A DC offset measurement is repeated 100 times, resulting in a series of values having an average of 257 mV and a standard deviation of 27 mV. In what range does the noiseless measured value lie for a 99.7% confidence? What is the accuracy of this measurement assuming the systematic offset is zero?

Solution:
Using Eq. (5.22) with $a = 3$, we can bound the noiseless measured value to lie in the range defined by

$$257\text{ mV} - 3 \times \frac{27\text{ mV}}{\sqrt{100}} \le E[V_{MEASURED}] \le 257\text{ mV} + 3 \times \frac{27\text{ mV}}{\sqrt{100}}$$

or

$$248.9\text{ mV} \le E[V_{MEASURED}] \le 265.1\text{ mV}$$

The accuracy of this measurement (where a measurement is the average of 100 voltage samples) would then be ±8.1 mV with a 99.7% confidence. Alternatively, if we repeat this measurement 1000 times, we can expect that 997 measured values (i.e., each measured value corresponding to 100 samples) will lie between 248.9 mV and 265.1 mV.

Exercises

5.6. A series of 100 measurements is made on the output of an op-amp circuit whereby the distribution was found to be Gaussian with mean value of 12.5 mV and a standard deviation of 10 mV. Write an expression for the pdf of these measurements?

ANS. $g(v) = \frac{1}{10^{-2}\sqrt{2\pi}} e^{-\frac{(v-12.5\times10^{-3})^2}{2(10^{-2})^2}}$, $\mu = 12.5$ mV; $\sigma = 10.0$ mV.

5.7. A series of 100 measurements is made on the output of an op-amp circuit whereby the distribution was found to be Gaussian with mean value of 12.5 mV and a standard deviation of 1 mV. If this experiment is repeated, write an expression for the pdf of the mean values of each of these experiments?

ANS. $f(v) = \frac{1}{(10^{-3})\sqrt{2\pi}} e^{-\frac{(v-12.5\times10^{-3})^2}{2(10^{-3})^2}}$, $\mu = 12.5$ mV; $\sigma = 1$ mV.

Exercises

5.8. A series of 100 measurements is made on the output of an op-amp circuit whereby the distribution was found to be Gaussian with mean value of 12.5 mV and a standard deviation of 1 mV. If this experiment is repeated and the mean value is compared to a reference level of 10 mV, what is the mean and standard deviation of the error distribution that results? Write an expression for the pdf of these errors?

ANS. $h(v) = \frac{1}{(10^{-3})\}\text{fin}} e$ - -2.5x10~3j 20o~3j

$\mu = 2.5$ mV; $a =$ 1 mV.

5.9. A series of 100 measurements is made on the output of an op amp circuit whereby the distribution was found to be Gaussian with mean value of 12.5 mV and a standard deviation of 1 mV. If this experiment and the mean value is compared to a reference value of 10 mV, in what range will the expected value of the error lie for a 99.7% conference interval.

ANS. -0.5 mV E[error] 5.5 mV

5.4 CALIBRATIONS AND CHECKERS

Measurement accuracy can be improved by eliminating the bias error p associated with a measurement process. This section will look at the several ways in which to remove this error so that the tester is performing with maximum accuracy at all times during its operation.

5.4.1 Traceability to Standards

Every tester and bench instrument must ultimately correlate to standards maintained by a central authority, such as the National Institute of Standards and Technology (NIST). In the United States, this government agency is responsible for maintaining the standards for pounds, gallons, inches, and electrical units such as volts, amperes, and ohms. The chain of correlation between the NIST and the tester's measurements involves a series of calibration steps that transfers the "golden" standards of the NIST to the tester's measurement instruments.

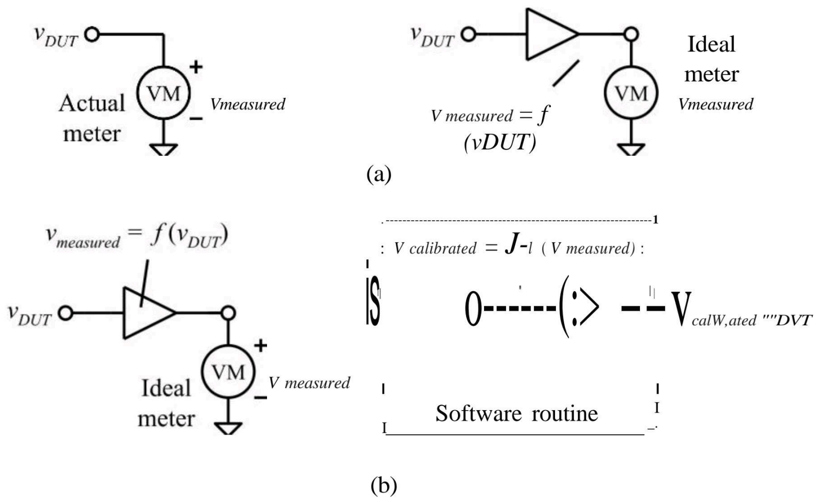
Many testers have a centralized standards reference, which is a thermally stabilized instrument in the tester mainframe. The standards reference is periodically replaced by a freshly calibrated reference source. The old one is sent back to a certified calibration laboratory, which recalibrates the reference so that it agrees with NIST standards. Similarly, bench instruments are periodically recalibrated so that they too are traceable to the NIST standards. By periodically refreshing the tester's traceability link to the NIST, all testers and bench instruments can be made to agree with one another.

5.4.2 Hardware Calibration

Hardware calibration is a process of physical "knob tweaking" that brings a piece of measurement instrumentation back into agreement with calibration standards. For instance, oscilloscope probes often include a small screw that can be used to nullify the overshoot in rapidly rising digital edges. This is one common example of hardware calibration.

One major problem with hardware calibration is that it is not a convenient process. It generally requires a manual adjustment of a screw or knob. Robotic screwdrivers might be employed to allow partial automation of the hardware calibration process. However, the use of robotics is an elaborate solution to the calibration problem. Full automation can be achieved using a simpler procedure known as *software calibration.*

Figure 5.5. [a) Modeling a voltmeter with an ideal voltmeter and a nonideal component in cascade. [b) Calibrating the nonideal effects using a software routine.



5.4.3 Software Calibration

Using software calibration, ATE testers are able to correct hardware errors without adjusting any physical knobs. The basic idea behind software calibration is to separate the instrument's ideal operation from its nonidealities. Then a *model* of the instrument's nonideal operation can be constructed, followed by a *correction* of the nonideal behavior using a mathematical routine written in software. Figure 5.5 illustrates this idea for a voltmeter.

In Figure 5.5a a "real" voltmeter is modeled as a cascade of two parts: (1) an ideal voltmeter and (2) a black box that relates the voltage across its input terminals v_{DUT} to the voltage that is measured by the ideal voltmeter, $v_{measured}$. This relationship can be expressed in more mathematical terms as

$$v_{measured} = f(v_{DUT}) \tag{5.27}$$

where $f(\cdot)$ indicates the functional relationship between $v_{measured}$ and v_{DUT}.

The true functional behavior $f(\cdot)$ is seldom known; thus one assumes a particular behavior or model, such as a first-order model given by

$$v_{measured} = G v_{DUT} + offset \tag{5.28}$$

where G and *offset* are the gain and offset of the voltmeter, respectively. These values must be determined from measured data. Subsequently, a mathematical procedure is written in software that performs the inverse mathematical operation

$$v_{calibrated} = f^{-1}(v_{measured}) \tag{5.29}$$

where $v_{calibrated}$ replaces v_{DUT} as an estimate of the true voltage that appears across the terminals of the voltmeter as depicted in Figure 5.5b. If $f(\cdot)$ is known precisely, then $v_{calibrated} = v_{DUT}$.

In order to establish an accurate model of an instrument, precise reference levels are necessary. The number of reference levels required to characterize the model fully will depend on its order-that is, the number of parameters used to describe the model. For the linear or firstorder model described, it has two parameters, G and *offset.* Hence, two reference levels will be required.

To avoid conflict with the meter's normal operation, relays are used to switch in these reference levels during the calibration phase. For example, the voltmeter in Figure 5.6 includes a pair of calibration relays, which can connect the input to two separate reference levels, V_{ref1} and V_{ref2}. During a system level calibration, the tester closes one relay and connects the voltmeter to V_{ref1} and measures the voltage, which we shall denote as $V_{measured1}$. Subsequently, this process is repeated for the second reference level V_{ref2} and the voltmeter provides a second reading, $V_{measured2}$.

Based on the assumed linear model for the voltmeter, we can write two equations in terms of two unknowns

$$
\begin{aligned}
V_{measured1} &= G V_{ref1} + offset \\
V_{measured2} &= G V_{ref2} + offset
\end{aligned}
\tag{5.30}
$$

Using linear algebra, the two model parameters can then be solved to be

$$
G = \frac{V_{measured2} - V_{measured1}}{V_{ref2} - V_{ref1}}
\tag{5.31}
$$

and

$$
offset = \frac{V_{measured1} V_{ref2} - V_{measured2} V_{ref1}}{V_{ref2} - V_{ref1}}
\tag{5.32}
$$

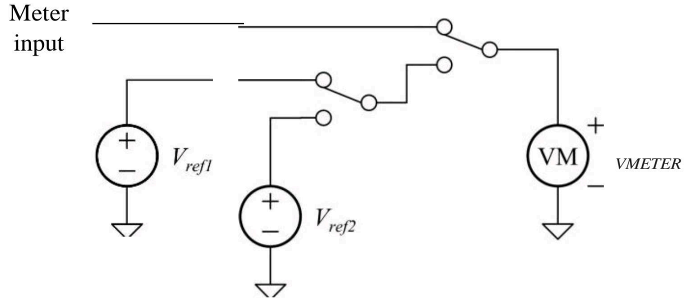
The parameters of the model, G and offset, are also known as *calibration factors,* or cal factors for short.

When subsequent DC measurements are performed, they are corrected using the stored calibration factors according to

$$
V_{calibrated} = \frac{V_{measured} - offset}{G}
\tag{5.33}
$$

This expression is found by isolating V_{DUT} on one side of the expression in Eq. (5.28) and replacing it by $V_{calibrated}$.

Figure 5.6. DC voltmeter gain and offset calibration paths.



Of course, this example is only for purposes of illustration. Most testers use much more elaborate calibration schemes to account for linearity errors and other nonideal behavior in the meter's ADC and associated circuits.. Also, the meter's input stage can be configured many ways, and each of these possible configurations needs a separate set of calibration factors. For example, if the input stage has 10 different input ranges, then each range setting requires a separate set of calibration factors. Fortunately for the test engineer, most instrument calibrations happen behind the scenes. The calibration factors are measured and stored automatically during the tester's periodic system calibration and checker process.

5.4.4 System Calibrations and Checkers

Testers are calibrated on a regular basis to maintain traceability of each instrument to the tester's calibration reference source. In addition to calibrations, software is also executed to verify the functionality of hardware and make sure it is production worthy. This software is called a *checker program,* or *checker* for short. Often calibrations and checkers are executed in the same program. If a checker fails, the repair and maintenance **(R&M)** staff replaces the failing tester module with a good one. After replacement, the new module must be completely recalibrated.

There are several types of calibrations and checkers. These include calibration reference source replacement, performance verification (PV), periodic system calibrations and checkers, instrument calibrations at load time, and focused calibrations. Calibration reference source replacement and recalibration was discussed in Section 5.4.1. A common replacement cycle time for calibration sources is once every six months.

Exercises

5.10. A meter reads 0.5 mV and 1.1 V when connected to two precision reference levels of 0 and 1 V, respectively. What are the offset and gain of this meter? Write the calibration equation for this meter.

ANS. 0.5 mV, 1.0995 V/V, $v_{calibrated} = (v_{measured} - 0.5\text{ mV})/1.0995$.

5.11. A meter is assumed characterized by a second-order equation of the form: $v_{measured} = offset + G_1 v_{calibrated} + G_2 v_{calibrated}^2$. How many precision DC reference levels are required to obtain the parameters of this second-order expression?

ANS. Three.

5.12. A meter is assumed characterized by a second-order equation of the form $v_{measured} = offset + G_1 v_{IN} + G_2 v_{IN}^2$. Write the calibration equation for this meter in terms of the unknown calibration factors.

ANS.

$$v_{calibrated} = \frac{-G_1 + \sqrt{G_1^2 + 4G_2 (v_{measured} - offset)}}{2G_2} \quad \text{or}$$

$$v_{calibrated} = \frac{-G_1 - \sqrt{G_1^2 + 4G_2 (v_{measured} - offset)}}{2G_2}$$

depending on the data conditions.

To verify that the tester is in compliance with all its published specifications, a more extensive process called *performance verification* may be performed. Although full performance verification is typically performed at the tester vendor's production floor, it is seldom performed on the production floor. By contrast, periodic system calibrations and checkers are performed on a regular basis in a production environment. These software calibration and checker programs verify that all the system hardware is production worthy.

Since tester instrumentation may drift slightly between system calibrations, the tester may also perform a series of fine-tuning calibrations each time a new test program is loaded. The extra calibrations can be limited to the subset of instruments used in a particular test program. This helps to minimize program load time. To maintain accuracy throughout the day, these calibrations may be repeated on a periodic basis after the program has been loaded. They may also be executed automatically if the tester temperature drifts by more than a few degrees.

Finally, focused calibrations are often required to achieve maximum accuracy and to compensate for nonidealities of DIB board components such as buffer amplifiers and filters. Unlike the ATE tester's built-in system calibrations, focused calibration and checker software is the responsibility of the test engineer. Focused calibrations fall into two categories: (1) focused instrument calibrations and (2) focused DIB calibrations and checkers.

5.4.5 Focused Instrument Calibrations

Testers typically contain a combination of slow, accurate instruments and fast instruments that may be less accurate. The accuracy of the faster instruments can be improved by periodically referencing them back to the slower more accurate instruments through a process called *focused calibration.* Focused calibration is not always necessary. However, it may be required if the test engineer needs higher accuracy than the instrument is able to provide using the built-in calibrations of the tester's operating system.

A simple example of focused instrument calibration is a DC source calibration. The DC sources in a tester are generally quite accurate, but occasionally they need to be set with minimal DC level error. A calibration routine that determines the error in a DC source's output level can be added to the first run of the test program. A high-accuracy DC voltmeter can be used to measure the actual output of the DC source. If the source is in error by 1 mV, for instance, then the requested voltage is reduced by 1 mV and the output is retested. It may take several iterations to achieve the desired value with an acceptable level of accuracy.

A similar approach can be extended to the generation of a sinusoidal signal requiring an accurate RMS value from an arbitrary waveform generator (AWG). A high-accuracy AC voltmeter is used to measure the RMS value from the AWG. The discrepancy between the measured value and the desired value is then used to adjust the programmed AWG signal level. The AWG output level will thus converge toward the desired RMS level as each iteration is executed.

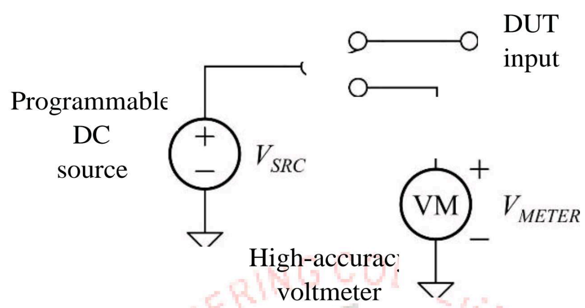
EXAMPLE 5.3

A 2.500-V signal is required from a DC source as shown in Figure 5.7. Describe a calibration procedure that can be used to ensure that 2.500 V ± 500 μV does indeed appear at the output of the DC source.

Solution:

The source is set to 2.500 V and a high-accuracy voltmeter is connected to the output of the source using a calibration path internal to the tester. Calibration path connections are made through one or more relays such as the ones in Figure 5.6. Assume the high-accuracy voltmeter reads 2.510 V from the source. The source is then reprogrammed to 2.500 V - 10 mV and the output is remeasured. If the second meter reading is 2.499 V, then the source is reprogrammed to 2.500 V - 10 mV + 1 mV and measured again. This process is repeated until the meter reads 2.500 V [plus or minus 500 μV). Once the exact programmed level is established, it is stored as a calibration factor [e.g., calibration factor= 2.500 V - 10 mV + 1 mV = 2.491 V). When the 2.500-V DC level is required during subsequent program executions, the 2.491 V calibration factor is used as the programmed level rather than 2.500 V. Test time is not wasted searching for the ideal level after the first calibration is performed. However, calibration factors may need to be regenerated every few hours to account for slow drifts in the DC source. This recalibration interval is dependent on the type of tester used.

Figure 5.7. DC source focused calibration.



Another application of focused instrument calibration is spectral leveling of the output of an AWG. An important application of AWGs is to provide a composite signal consisting of *N* sine waves or *tones* all having equal amplitude at various frequencies and arbitrary phase. Such waveforms are in a class of signals commonly referred to as *multitone* signals. Mathematically, a multitone signal *y(t)* can be written as

$$y(t) = A_0 + A_1 \sin(2\pi f_1 t + \phi_1) + \cdots + A_N \sin(2\pi f_N t + \phi_N) = A_0 + \sum_{k=1}^{N} A_k \sin(2\pi f_k t + \phi_k) \quad \textbf{(5.34)}$$

where A_k, f_k, and ϕ_k denote the amplitude, frequency, and phase, respectively, of the kth tone. A multitone signal can be viewed in either the time domain or in the frequency domain. Timedomain views are analogous to oscilloscope traces, while frequency-domain views are analogous to spectrum analyzer plots. The frequency-domain graph of a multitone signal contains a series of vertical lines corresponding to each tone frequency and whose length* represents the root-mean-square **(RMS)** amplitude of the corresponding tone. Each line is referred to as a *spectral line.* Figure 5.8 illustrates the time and frequency plots of a composite signal consisting of three tones of frequencies 1, 2.5, and 4.1 kHz, all having an RMS amplitude of 2 V. Of course, the peak amplitude of each sinusoid in the multitone is simply $\sqrt{2} \times 2$ or 2.82 V, so we could just as easily plot these values as peak amplitudes rather than **RMS.** This book will consistently display frequency-domain plots using RMS amplitudes.

*Spectral density plots are commonly defined in engineering textbooks with the length of the spectral line representing one-half the amplitude of a tone. In most test engineering work, including spectrum analyzer displays, it is more common to find this length defined as an RMS quantity.

Figure 5.8. Time-domain and frequency-domain views of a three-tone multitone.

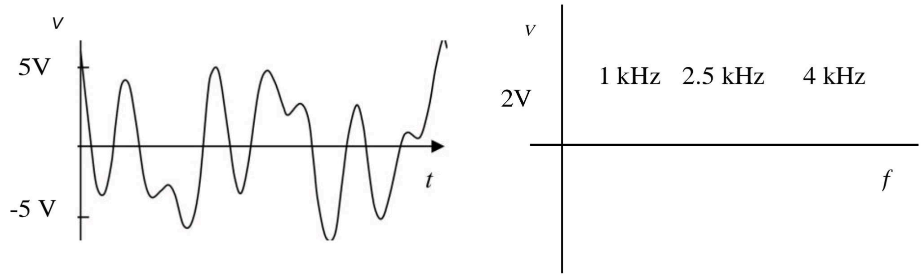
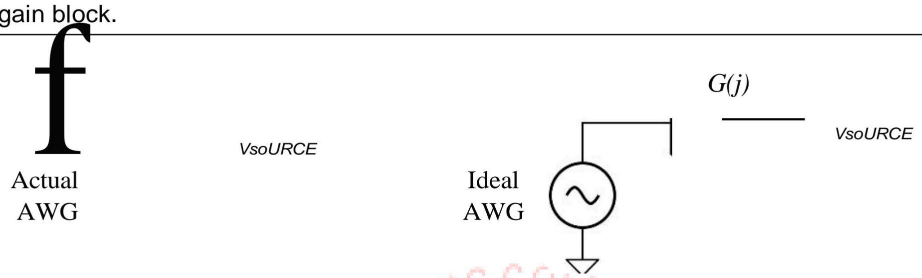


Figure 5.9. Modeling an AWG as a cascaded combination of an ideal source and frequency-dependent gain block.



The AWG produces its output signal by passing the output of a DAC through a low-pass antiimaging filter. Due to its frequency behavior, the filter will not have a perfectly flat magnitude response. The DAC may also introduce frequency-dependent errors. Thus the amplitudes of the individual tones may be offset from their desired levels. We can therefore model this AWG multitone situation as illustrated in Figure 5.9. The model consists of an ideal source connected in cascade with a linear block whose gain or magnitude response is described by $G(f)$, where f is the frequency expressed in hertz. To correct for the gain change with frequency, the amplitude of each tone from the AWG is measured individually using a high-accuracy AC voltmeter. The ratio between the actual output and the requested output corresponds to $G(f)$ at that frequency. This gain can then be stored as a calibration factor that can subsequently be retrieved to correct the amplitude error at that frequency. The calibration process is repeated for each tone in the multitone signal. The composite signal can then be generated with corrected amplitudes by dividing the previous requested amplitude at each frequency by the corresponding AWG gain calibration factor. Because the calibration process equalizes the amplitudes of each tone, the process is called multitone leveling.

As testers continue to evolve and improve, it may become increasingly unnecessary for the test engineer to perform focused calibrations of the tester instruments. Focused calibrations were once necessary on almost all tests in a test program. Today, they can sometimes be omitted with little degradation in accuracy. Nevertheless, the test engineer must evaluate the need for focused calibrations on each test. Even if calibrations become unnecessary in the future, the test engineer should still understand the methodology so that test programs on older equipment can be comprehended.

Calibration of circuits on the DIB, on the other hand, will probably always be required. The tester vendor has no way to predict what kind of buffer amplifiers and other circuits will be placed on the DIB board. The tester operating system will never be able to provide automatic calibration of these circuits. The test engineer is fully responsible for understanding the calibration requirements of all DIB circuits.

Exercises

5.13. A DC source is assumed characterized by a third-order equation of the form $v_{MEASURED} = 0.005 + v_{PROGRAMMED} - 0.003\, v;_{ROGRAMMED}$ and is required to generate a DC level of 2.6 V. However, when programmed to produce this level, only 2.552 V is measured. Using iteration, determine a value of the programmed source voltage that will establish a measured voltage of 2.6 V to within a ± 1-mV accuracy.

ANS. 2.651 V.

5.14. An AWG has a gain response described by ($1+ (1/_{,0'})^{2}$ r and is to generate three tones at frequencies of 1, 2, and 3 kHz. What are the calibration factors?

ANS. 0.707, 0.447, and 0.316.

5.4.6 Focused DIB Circuit Calibrations

Often circuits are added to a DIB board to improve the accuracy of a particular test or to buffer the weak output of a device before sending it to the tester electronics. As the signal-conditioning DIB circuitry is added in cascade with the test instrument, a model of the test setup is identical to that given in Figure 5.5a. The only difference is that functional block $v_{measured} = f\left(v_{DUT}\right)$ includes both the meter and the DIB's behavior. As a result, the focused instrument calibrations of Section 5.4.3 can be used with no modifications. Conversely, the meter may already have been calibrated so that the functional block/(·) covers the DIB circuitry only. One must keep track of the extent of the calibration to avoid any double counting.

EXAMPLE 5.4

The op-amp circuit in Figure 5.10 has been added to a 01B board to buffer the output of a OUT. The buffer will be used to condition the DC signal from the OUT before sending it to a calibrated DC voltmeter resident in the tester. If the output is not buffered, then we may find that the OUT breaks into oscillations as a result of the stray capacitance arising along the lengthy signal path leading from the OUT to the tester. The buffer prevents these oscillations by substantially reducing stray capacitance at the OUT output. In order to perform an accurate measurement, the behavior of the buffer must be accounted for. Outline the steps to perform a focused DC calibration on the op-amp buffer stage.

Solution:

To perform a DC calibration of the output buffer amplifier, it is necessary to assume a model for the op-amp buffer stage. It is reasonable to assume that the buffer is fairly linear over a wide range of signal levels, so that the following linear model can be used:

$$v_{measured} = G v_{OUT} + \text{offset}$$

Subsequently, following the same procedure as outlined in Section 5.4.3, a pair of known voltages are applied to the input of the buffer from source SRC1 via the relay connection and the output of the buffer is measured with a voltmeter. This temporary connection is called a calibration path. As an example, let SRC1 force 2 V and assume that an output voltage of 2.023 Vis measured using the voltmeter. Next the input is dropped to 1 V, resulting in an output voltage of 1.012 V. Using Eq. [5.31], we find the buffer has gain given by

$$G = \frac{2.023\,V - 1.012\,V}{2\,V - 1\,V} = 1.011\ V/V$$

and the offset is found from Eq. [5.32) to be

$$offset = \frac{1.012\,V \cdot 2\,V - 2.023\,V \times 1\,V}{2\,V - 1\,V} = 1\ mV$$

Hence, the OUT output v_{OUT} and the voltmeter value $v_{measured}$ are related according to

$$v_{measured} = 1.011\ V/V \times v_{OUT} + 0.001\ V$$

The goal of the focused DC calibration procedure is to find an expression that relates the OUT output in terms of the measured value. Hence, by rearranging the expression and replacing $v_{calibrated}$ for v_{OUT}, we obtain

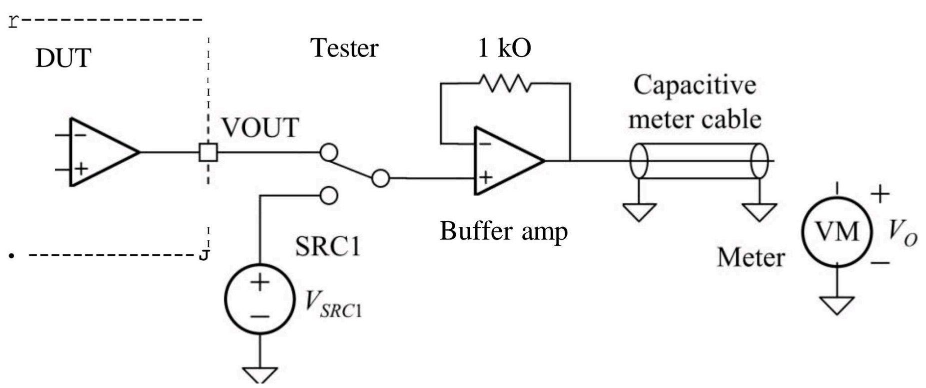
$$v_{calibrated} = \frac{v_{measured} - 0.001\ V}{1.011\ V/V}$$

For example, if the voltmeter reads 1.732 V, the actual voltage appearing at its terminals is actually

$$v_{calibrated} = \frac{1.732V - 0.001V}{1.011\ V/V} = 1.712V$$

If the original uncalibrated answer had been used, there would have been a 20-mV error! This example shows why focused OUT calibrations are so important to accurate measurements.

Figure 5.10. DC calibration for op-amp buffer circuit.



When buffer amplifiers are used to assist the measurement of AC signals, a similar calibration process must be performed on each frequency that is to be measured. Like the AWG calibration example, the buffer amplifier also has a nonideal frequency response and will affect the reading of the meter. Its gain variation, together with the meter's frequency response, must be measured at each frequency used in the test during a calibration run of the test program. Assuming that the meter has already been calibrated, the frequency response behavior of the DIB circuitry must be correctly accounted for. This is achieved by measuring the gain in the DIB's signal path at each specific test frequency. Once found, it is stored as a calibration factor. If additional circuits such as filters, ADCs, and so on, are added on the DIB board and used under multiple configurations, then each unique signal path must be individually calibrated.

5.5 TESTER SPECIFICATIONS

The test engineer should exercise diligence when evaluating tester instrument specifications. It can be difficult to determine whether or not a particular tester instrument is capable of making a particular measurement with an acceptable level of accuracy. The tester specifications usually do not include the effects of uncertainty caused by instrument repeatability limitations. All the specification conditions must be examined carefully. Consider the following DC meter example.

A DC meter consisting of an analog-to-digital converter and a programmable gain amplifier (PGA) is shown in Figure 5.11. The programmable gain stage is used to set the range of the meter so that it can measure small signals as well as large ones. Small signals are measured with the highest gain setting of the PGA, while large signals are measured with the lowest gain setting. This ranging process effectively changes the resolution of the ADC so that its quantization error is minimized.

Calibration software in the tester compensates for the different PGA gain settings so that the digital output of the meter's ADC can be converted into an accurate voltage reading. The calibration software also compensates for linearity errors in the ADC and offsets in the PGA and ADC. Fortunately, the test engineer does not have to worry about these calibrations because they happen automatically.

Table 5.1 shows an example of a specification for a fictitious DC meter, the DVMIO0. This meter has five different input ranges, which can be programmed in software. The different ranges allow small voltages to be measured with greater accuracy than large voltages. The accuracy is specified as a percentage of the measured value, but there is an accuracy limit of 1 mV for the lower ranges and 2.5 mV for the higher ranges.

Figure 5.11. Simplified DC voltmeter with input ranging amplifier.

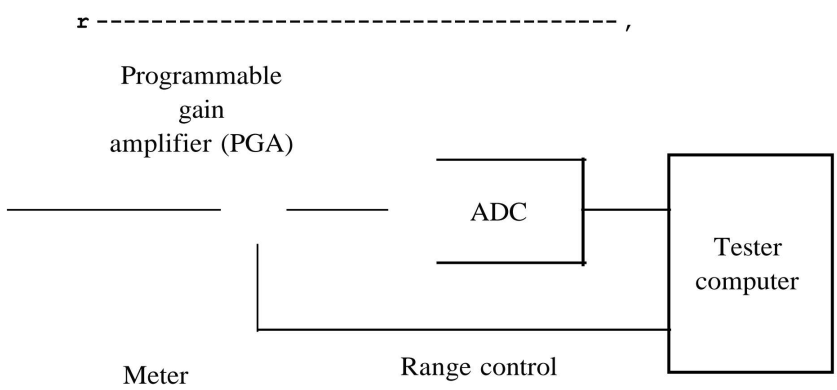


Table 5.1. DVM100 DC Voltmeter Specifications

Range	Resolution	Accuracy(% of Measurement)
±0.5 V	15.25 μV	±0.05 % or 1 mV [whichever is greater]
±1 V	30.5 μV	±0.05 % or 1 mV
±2 V	61.0 μV	±0.05 % or 1 mV
±5 V	152.5 μV	±0.10 % or 2.5 mV
±10 V	305.2 mV	±0.10 % or 2.5 mV

Note: All specs apply with the measurement filter enabled.

This accuracy specification probably assumes that the measurement is made 100 or more times and averaged. For a single nonaveraged measurement, there may also be a repeatability error to consider. It is not clear from the table above what assumptions are made about averaging. The test engineer should make sure that all assumptions are understood before relying on the accuracy numbers.

EXAMPLE 5.5

A OUT output is expected to be 100 mV. Our fictitious DC voltmeter, the DVM 100, is set to the 0.5-V range to achieve the optimum resolution and accuracy. The reading from the meter [with the meter's input filter enabled) is 102.3 mV. Calculate the accuracy of this reading [excluding possible repeatability errors). What range of outputs could actually exist at the OUT output with this reading?

Solution:

The measurement error would be equal to ±0.05% of 100 mV, or 50 μV, but the specification has a lower limit of 1 mV. The accuracy is therefore ±1 mV. Based on the single reading of 102.3 mV, the actual voltage at the OUT output could be anywhere between 101.3 and 103.3 mV.

In addition to the ranging hardware, the meter also has a low-pass filter in series with its input. The filter can be bypassed or enabled, depending on the measurement requirements. Repeatability is enhanced when the low-pass filter is enabled, since the filter reduces electrical noise in the input signal. Without this filter the accuracy would be degraded by nonrepeatability. The filter undoubtedly adds settling time to the measurement, since all low-pass filters require time to stabilize to a final DC value. The test engineer must often choose between slow, repeatable measurements and fast measurements with less repeatability.

It may be possible to empirically determine through experimentation that this DC voltmeter has adequate resolution and accuracy to make a DC offset measurement with less than 100 μV of error. However, since this level of accuracy is far better than the instrument's ±1-mV specifications, the instrument should probably not be trusted to make such a measurement in production. The accuracy might hold up for 100 days and then drift toward the specification limits of 1 mV on day 101.

Another possible scenario is that multiple testers may be used that do not all have 100-μV performance. Tester companies are often conservative in their published specifications, meaning that the instruments are often better than their specified accuracy limits. This is not a license to use the instruments to more demanding specifications. It is much safer to use the specifications as printed, since the vendor will not take any responsibility for use of instruments beyond their official specifications.

Sometimes the engineer may have to design front-end circuitry such as PGAs and filters onto the DIB board itself. The DIB circuits might be needed if the front-end circuitry of the meter is inadequate for a high-accuracy measurement. Front-end circuits may also be added if the signal from the DUT cannot be delivered cleanly through the signal paths to the tester instruments. Very high-impedance DUT signals might be susceptible to externally coupled noise, for example. Such signals might benefit from local buffering and amplification before passing to the tester instrument. The test engineer must calibrate any such buffering or filtering circuits using a focused DIB calibration.

Exercises

5.15.	A voltmeter is specified to have an accuracy of ±1% of programmed range. If a DC level is measured on a ±1 V range and appears on the meter as 0.5 V, what are the minimum and maximum DC levels that might have been present at the meter's input during this measurement?	ANS. 0.5 V ±10 mV [i.e., the input could lie anywhere between 490 and 510 mV).

5.6 REDUCING MEASUREMENT ERROR WITH GREATER MEASUREMENT TIME

In this section we shall look at several commonly used filtering techniques for extracting the mean value of a measurement process. As we shall see, filtering will improve measurement accuracy at the expense of longer test time.

5.6.1 Analog Filtering

Analog filters are often used in tester hardware to remove unwanted signal components before measurement. A DC voltmeter may include a low-pass filter as part of its front end. The purpose of the filter is to remove all but the lowest-frequency components. It acts as a hardware averaging circuit to improve the repeatability of the measurement. More effective filtering is achieved using a filter with a low cutoff frequency, since a lower cutoff frequency excludes more electrical noise. Consequently, a lower frequency cutoff corresponds to better repeatability in the final measurement. For instance, if the low-pass filter has cutoff frequency *mb* and the variation of a measurement has a bandwidth of *mnur* (assumed to be much less than *mb)* with an RMS value of *V nur,* then we can compute the value of the RMS noise that passes through the filter according to

$$V n_0 = V nur \sqrt{\frac{\omega_b}{\omega_{DUT}}} \quad \text{V} \tag{5.35}$$

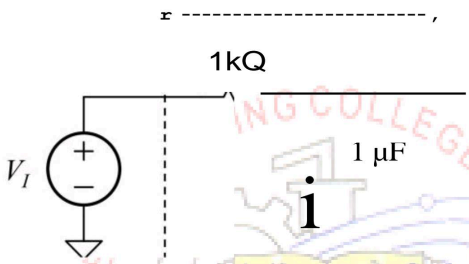
The above expression illustrates the noise reduction gained by filtering the output. The smaller the ratio *mb lmDuT'* the greater the noise reduction. Other types of filtering circuits can be placed on the DIB board when needed. For example, a very narrow bandpass filter may be placed on the DIB board to clean up noise components in a sine wave generated by the tester. The filter allows a much more ideal sine wave to the input of the DUT than the tester would otherwise be able to produce.

An important drawback to filtering a signal prior to measurement is the additional time required for the filter to settle to its steady-state output. The settling time is inversely proportional to the filter cutoff frequency. Thus, there is an inherent tradeoff between repeatability and test time. The following example will quantify this tradeoff for a first-order system.

EXAMPLE 5.6

The simple RC low-pass circuit shown in Figure 5.12 is used to filter the output of a OUT containing a noisy DC signal. For a particular measurement, the signal component is assumed to change from Oto 1 V, instantaneously. How long does it take the filter to settle to within 1% of its final value? By what factor does the settling time increase when the filter's 3-dB bandwidth is decreased by a factor of 10?

Figure 5.12. RC low-pass filter.



Solution:

From the theory of first-order networks, the step response of the circuit starting from rest (i.e., $v_{,}$ = 0) is

$$v_o(t) = S\left(1 - e^{-t/\tau}\right) \tag{5.36}$$

where S = 1 Vis the magnitude of the step and "c $=RC=10^{-3}$ s. Moreover, the 3-dB bandwidth *rob* (expressed in rad/s) of a first-order network is 1/RC, so we can rewrite the above expression as

$$v_o(t)=S\ (1\text{-e}^{-OJ_b\ t)} \tag{5.37}$$

Clearly, the time $t = t_s$ the output reaches an arbitrary output level of V_o is then

$$\frac{\ln\left(\frac{S-V_o}{S}\right)}{\omega_b} \tag{5.38}$$

Furthermore, we recognize that $(S - V_0)/S$ is the settling error ε or the accuracy of the measurement, so we can rewrite Eq. [5.38] as

$$t_S = -\frac{\ln(\varepsilon)}{\omega_b} \tag{5.39}$$

Hence, the time it takes to reach within 1% of 1 V, or 0.99 V, is 4.6 ms. Since settling time and 3-dB bandwidth are inversely related according to Eq. [5.39], a tenfold decrease in bandwidth leads to a tenfold increase in settling time. Specifically, the settling time becomes 46 ms.

Exercises

5.16. What is the 3-dB bandwidth of the *RC* circuit of Figure 5.12 expressed in hertz, when R = 1 kΩ and C = 2.2 nF?

ANS. 72.34 kHz.

5.17. How long does it take a first-order *RC* low-pass circuit with R = 1 kΩ and C = 2.2 nF to settle to 5% of its final value?

ANS. 6.6 μs.

5.18. By what factor should the bandwidth of an *RC* low-pass filter be decreased in order to reduce the variation in a DC measurement from 250 μV RMS to 100 μV RMS. By what factor does the settling time increase.

ANS. The bandwidth should be decreased by 6.25 [$= 2.5^2$]. Settling time increases by 6.25.

5.19. The variation in the output signal of a OUT is 1 mV RMS. Assume that the DUT's output follows a first-order frequency response and has a 3-dB bandwidth of 100 Hz. Estimate the output noise voltage spectral density.

ANS. 6.37×10^{-9} V^2/Hz.

5.20. The variation in the output RMS signal of a OUT is 1 mV, but it needs to be reduced to a level closer to 500 μV. What filter bandwidth is required to achieve this level of repeatability? Assume that the DUT's output follows a first-order frequency response and has a 3-dB bandwidth of 1000 Hz.

ANS. 250 Hz.

5.21. The output of a OUT has an uncertainity of 10 mV. How many samples should be combined in order to reduce the uncertainity to 100 μV?

ANS. 10,000.

5.6.2 Averaging

Averaging defined by the expression $\frac{1}{N}\sum_{k=1}^{N} x[k]$ is a specific form of discrete-time filtering. Averaging can be used to improve the repeatability of a measurement. For example, we can average the following series of nine voltage measurements and obtain an average of 102 mV.

101 mV, 103 mV, 102 mV, 101 mV, 102 mV, 103 mV, 103 mV, 101 mV, 102 mV

There is a good chance that a second series of nine unique measurements will again result in something close to 102 mV. If the length of the series is increased, the answer will become more repeatable and reliable. But there is a point of diminishing returns. To reduce the effect of noise on the voltage measurement by a factor of two, one has to take four times as many readings and average them. At some point, it becomes prohibitively expensive (i.e., from the point of view of test time) to improve repeatability. In general, if the RMS variation in a measurement is again denoted *V vuT,* then after averaging the measurement *N* time, the RMS value of the resulting averaged value will be

$$V_{n_o} = \frac{V_{DUT}}{\sqrt{N}} \quad \text{V} \tag{5.40}$$

Here we see the output noise voltage reduces the input noise before averaging by the factor $\sqrt{N}$. Hence, to reduce the noise RMS voltage by a factor of two requires an increase in the sequence length, *N,* by a factor of four.

AC measurements can also be averaged to improve repeatability. A series of sine wave signal level measurements can be averaged to achieve better repeatability. However, one should not try to average readings in decibels. If a series of measurements is expressed in decibels, they should first be converted to linear form using the equation *V* = *lQdBIW* before applying averaging. Normally, the voltage or gain measurements are available before they are converted to decibels in the first place; thus the conversion from decibels to linear units or ratios is not necessary. Once the average voltage level is calculated, it can be converted to decibels using the equation $\text{dB} = 20\log_{10}(\text{V})$. To understand why we should not perform averaging on decibels, consider the sequence 0, -20, -40 dBV. The average of these values is -20 dBV. However, the actual voltages are 1 V, 100 mV, and 10 mV. Thus the correct average value is (1 V + 0.1 V + 0.01 V) / 3 = 37 mV, or-8.64 dBV.

5.7 GUARDBANDS

Guardbanding is an important technique for dealing with the uncertainty of each measurement. If a particular measurement is known to be accurate and repeatable with a worst-case uncertainty of ±£, then the final test limits should be tightened from the data sheet specification limits by £ to make sure that no bad devices are shipped to the customer. In other words,

guardbanded upper test limit = upper specification limit - £
guardbanded lower test limit = lower specification limit +£ **(5.41)**

So, for example, if the data sheet limit for the offset on a buffer output is -100 mV minimum and 100 mV maximum, and an uncertainty of ±10 mV exists in the measurement, the test program limits should be set to -90 mV minimum and 90 mV maximum. This way, if the device output is 101 mV and the error in its measurement is -10 mV, the resulting reading of 91 mV will cause a failure as required. Of course, a reading of 91 mV may also represent a device with an 81-mV output and a +10-mV measurement error.

In such cases, guardbanding has the unfortunate effect of disqualifying good devices. Ideally, we would like all guardbands to be set to O so that no good devices will be discarded. To minimize the guardbands, we must improve the repeatability and accuracy of each test, but this typically requires longer test times. There is a balance to be struck between repeatability and the number of good devices rejected. At some point, the added test time cost of a more repeatable measurement

Table 5.2. OUT Output and Measured Values

DUT Output (mV)	Measured Value (mV)
105	101
101	107
98	102
96	95
86	92
72	78

outweighs the cost of discarding a few good devices. This tradeoff is illustrated in Figure 5.13 on the histogram of some arbitrary offset voltage test data for two different-sized guardbands. With larger guardbands, the region of acceptability is reduced; hence fewer good devices will be shipped.

EXAMPLE 5.7

Table 5.2 lists a set of output values from a OUT together with their measured values. It is assumed that the upper specification limit is 100 mV and the measurement uncertainty is ±6 mV. How many good devices are rejected because of the measurement error? How many good devices are rejected if the measurement uncertainty is increased to ±10 mV?

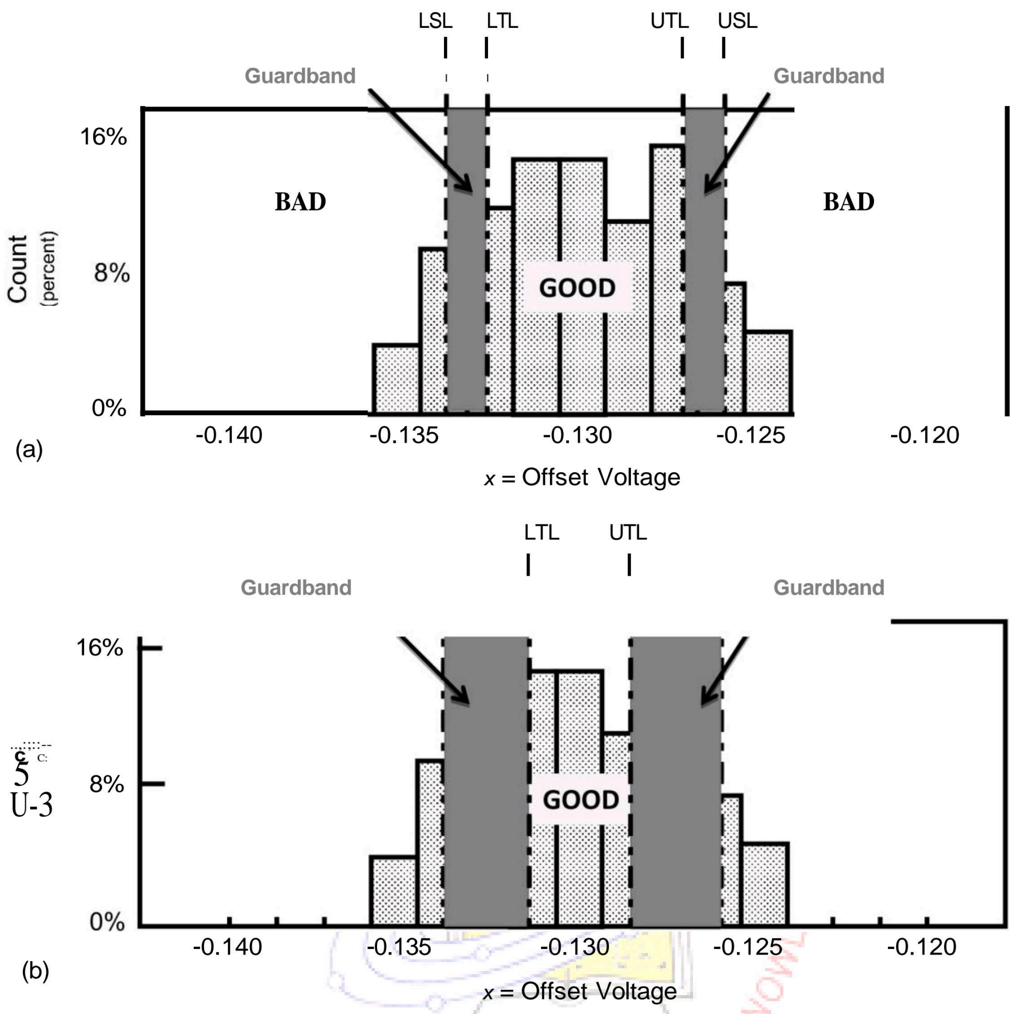
Solution:

From the OUT output column on the left, four devices are below the upper specification limit of 100 mV and should be accepted. The other two should be rejected. Now with a measurement uncertainty of ±6 mV, according to Eq. (5.41) the guardbanded upper test limit is 94 mV. With the revised test limit, only two devices are acceptable. The others are all rejected. Hence, two otherwise good devices are disqualified.

If the measurement uncertainty increases to ±10 mV, then the guardbanded upper test limit becomes 90 mV. Five devices are rejected and only one is accepted. Consequently, three otherwise good devices are disqualified.

In practice, we need to set E equal to 3 to 6 times the standard deviation of the measurement to account for measurement variability. A diagram illustrating the impact of shifting the test limits away from the specification limits on the probability density is provided in Figure 5.14. This diagram shows a marginal device with an average (true) reading equal to the upper specification limit. The upper and lower specification limits (USL and LSL, respectively) have each been tightened by E = 3cr. The tightened upper and lower test limits (UTL and LTL, respectively) reject marginal devices such as this, regardless of the magnitude of the measurement error. A more stringent guardband value of £ = *6cr* gives us an extremely low probability of passing a defective device, but this is sometimes too large a guardband to allow a manufacturable yield.

Figure 5.13. (a) Guardbanding the specification limits. (b) Illustrating the implications of large guardbands on the region of acceptability.



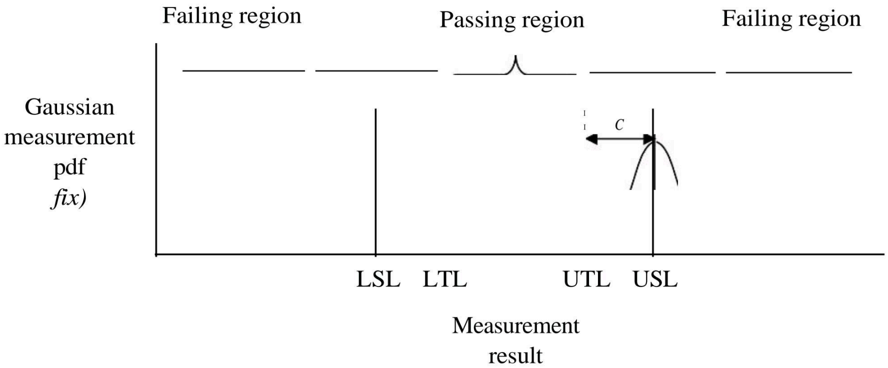
EXAMPLE 5.8

A DC offset measurement is repeated many times, resulting in a series of values having an average of 257 mV. The measurements exhibit a standard deviation of 27 mV. If our specification limits are 250 ±50 mV, where would we have to set our 6cr guardbanded upper and lower test limits?

Solution:

The value of cr is equal to 27 mV; thus the width of the 6cr guardbands would have to be equal to 162 mV. The upper test limit would be 300 mV - 162 mV, and the lower test limit would be 200 mV + 162 mV. Clearly, there is a problem with the repeatability of this test, since the lower guardbanded test limit is higher than the upper guardbanded test limit! Averaging would have to be used to reduce the standard deviation.

Figure 5.14. Guardbanded measurement with Gaussian distribution.



If a device is well-designed and a particular measurement is sufficiently repeatable, then there will be few failures resulting from that measurement. But if the distribution of measurements from a production lot is skewed so that the average measurement is close to one of the test limits, then production yields are likely to fall. In other words, more good devices will fall within the guardband region and be disqualified. Obviously, a measurement with poor accuracy or poor repeatability will just exacerbate the problem.

The only way the test engineer can minimize the required guardbands is to improve the repeatability and accuracy of the test, but this requires longer test times. At some point, the test time cost of a more repeatable measurement outweighs the cost of throwing away a few good devices. Thus there are inherent tradeoffs between repeatability, test time, guardbands, and production yield.

The standard deviation of a test result calculated as the average of N values from a statistical population is given by

$$\sigma_{ave} = \frac{\sigma}{\sqrt{N}} \tag{5.42}$$

So, for example, if we want to reduce the value of a measurement's standard deviation o- by a factor of two, we have to average a measurement four times. This gives rise to an unfortunate exponential tradeoff between test time and repeatability.

We can use Gaussian statistical analysis to predict the effects of nonrepeatability on yield. This allows us to make our measurements repeatable enough to give acceptable yield without wasting time making measurements that are *too* repeatable. It also allows us to recognize the situations where the average device performance or tester performance is simply too close to failure for economical production.

EXAMPLE 5.9

How many times would we have to average the DC measurement in Example 5.8 to achieve 6<3 guardbands of 10 mV? If each measurement takes 5 ms, what would be the total test time for the averaged measurement?

Solution:

The value of *a-ave* must be equal to 10 mV divided by 6 to achieve 60- guardbands. Rearranging Eq. [5.42]. we see that N must be equal to

$$N=\left(\frac{\sigma}{\sigma_{ave}}\right)^2=\left(\frac{27\text{ mV}}{10\text{ mV}/6}\right)^2=262\text{ measurements}$$

The total test time would be equal to 262 times 5 ms, or 1.31 s. This is clearly unacceptable for production testing of a DC offset. The 27-mV standard deviation must be reduced through an improvement in the 01B hardware or the OUT design.

Above we stated that the guardbands should be selected to be between 3 and 6 standard deviations of the measurement. Here we recast this statement in terms of the desired defect level. Consider the situation depicted in Figure 5.14 for a marginal device. The probability that a bad part will have a measured value that is less than UTL is given by

P(X<VTL)=<l>(:"J (5.43)

If N devices are produced, the defect level in ppm as defined by Eq. (5.3) can be written as

DL [ppm]= #es apes x 106 =<l>(:" J x 106 (5.44)

Rearranging Eq. (5.44) and isolating the guardband term we find

$$\varepsilon=\sigma_n\times\left|\Phi^{-1}\left(\frac{DL}{10^6}\right)\right| \tag{5.45}$$

The upper and lower test limits can then be found Eq. (5.41) above.

EXAMPLE 5.10

A DC offset test is performed on a OUT with lower and upper specification limits of -5 mV and 5 mV, respectively. The expected RMS level of the noise present during the test is 1 mV. If a defect level of less than 200 ppm is required, what should be the test limits?

Solution:

According to Eq. (5.45). the guardband is

$$\varepsilon = 10^{-3} \times \left| \Phi^{-1}\left(\frac{200}{10^6}\right) \right| = 35.40 \text{ mV}$$

then from Eq. (5-41). we find

LTL=-5+£=-1.45 mV and UTL = 5 - £ = 1.45 mV

Exercises

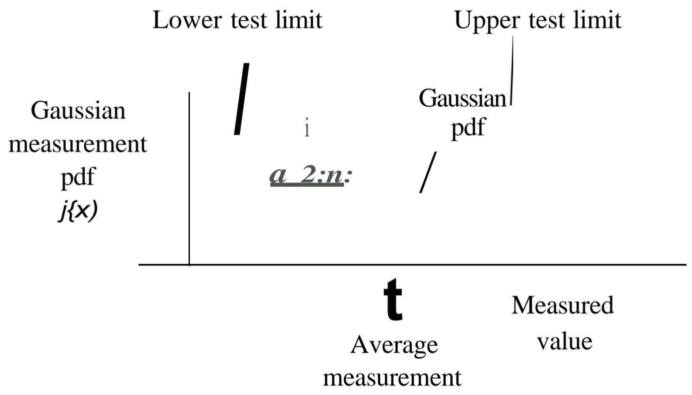
5.22.	A series of AC RMS measurements reveal an average value of 1.25 V and a standard deviation of 35 mV. If our specification limits were 1.2 V plus or minus 150 mV, where would we have to set our $3a$ guardbanded upper and lower test limits? If 6aguardbands are desired, how many times would we have to average the measurement to achieve guardbands of 40 mV?	ANS. $3a$ guardbanded test limits are: 1.15 and 1.245 V. N = 28.
5.23.	A device is expected to exhibit a worst-case offset voltage of ±50 mV and is to be measured using a voltmeter having an accuracy of only ±5 mV. Where should the guardbanded test limits be set?	ANS. ±45 mV.
5.24.	The guardband of a particular measurement is 10 mV and the test limit is set to ±25 mV. What are the original device specification limits?	ANS. ±35 mV.
5.25.	The following lists a set of output voltage values from a group of DUTs together with their measured values: {(2.3, 2.1l, [2.1, 1.6]. [2.2, 2.1]. [1.9, 1.6]. [1.8, 1.7]. [1.7, 2.1]. (1.5, 2.0)}. If the upper specification limit is 2.0 V and the measurement uncertainty is ±0.5 V, how many good devices are rejected due to the measurement error?	ANS. Four devices (all good devices are rejected by the 1.5-V guardbanded upper test limit).

5.8 EFFECTS OF MEASUREMENT VARIABILITY ON TEST YIELD

Consider the case of a measurement result having measurement variability caused by additive Gaussian noise. This test has a lower test limit (LTL) and an upper test limit (UTL). If the true measurement result is exactly between the two test limits, and the repeatability error never exceeds ±1/2 (UTL - LTL), then the test will always produce a passing result. The repeatability error never gets large enough to push the total measurement result across either of the test limits. This situation is depicted in Figure 5.15 where the pdf plot is shown.

On the other hand, if the average measurement is exactly equal to either the LTL or the UTL, then the test results will be unstable. Even a tiny amount of repeatability error will cause the test to randomly toggle between a passing and failing result when the test program is repeatedly executed. Assuming the statistical distribution of the repeatability errors is symmetrical, as in the

Figure 5.15. Probability density plot for measurement result between two test limits.



case of the Gaussian pdf, the test will produce an equal number of failures and passing results. This is illustrated by the pdf diagram shown in Figure 5.16. The area under the pdf is equally split between the passing region and the failing region; so we would expect 50% of the test results to pass and 50% to fail.

For measurements whose average value is close to but not equal to either test limit, the analysis gets a little more complicated. Consider an average measurement μ that is 8_1 units below the upper test limit as shown in Figure 5.17.

Any time the repeatability error exceeds 8_1 the test will fail. In effect, the measurement noise causes an erroneous failure. The probability that the measurement error will *not* exceed 8_1 and cause a failure is equal to the area underneath the portion of the pdf that is less than the UTL. This area is equal to the integral of the pdf from minus infinity to the UTL of the measurement results. In other words, the probability that a measurement will not fail the upper test limit as adopted from Eq. 5.19 is

$$P\left(X < \text{UTL}\right) = \Phi\left(\frac{\text{UTL} - \mu}{\sigma}\right) \quad \textbf{(5.46)}$$

Conversely, the probability of a failing result due to the upper test limit is

$$P\left(\text{UTL} < X\right) = 1 - \Phi\left(\frac{\text{UTL} - \mu}{\sigma}\right) \quad \textbf{(5.47)}$$

Similar equations can be written for the lower test limit.

If the distribution of measurement values becomes very large relative to the test limits, then we have to consider the area in both failing regions as shown in Figure 5.18. Clearly, if the true measurement result μ is near either test limit, or if the standard deviation cr is large, the test program has a much higher chance of rejecting a good DUT.

Considering both UTL failures and LTL failures, the probability of a passing result given this type of measurement repeatability according to Eq. (5.19) is

$$P\left(\text{LTL} < X < \text{UTL}\right) = \Phi(\text{U} - \mu) - \Phi(\text{LT} - \mu) \quad \textbf{(5.48)}$$

Figure 5.16. Probability density plot for nonrepeatable measurement centered at the UTL.

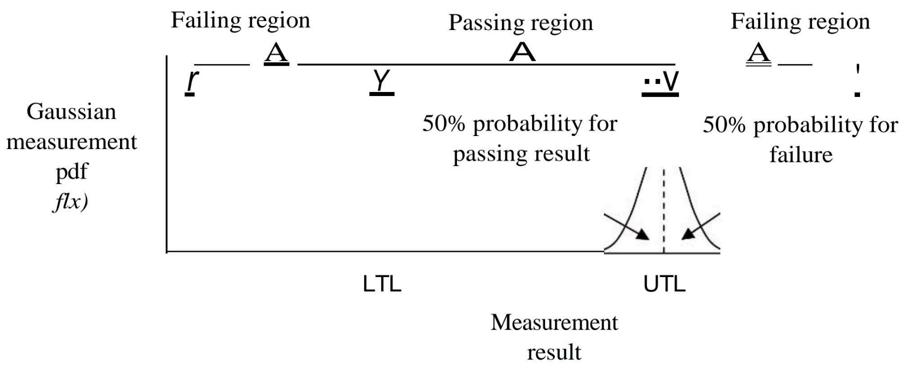
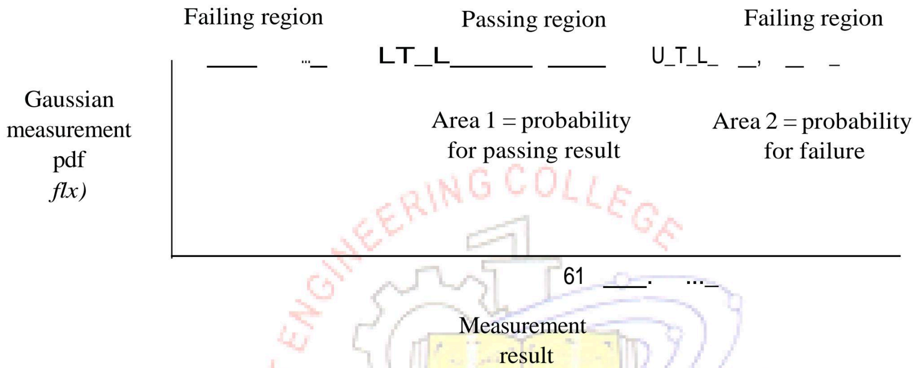


Figure 5.17. Probability density plot for average reading, μ, slightly below UTL by δ_1.



The probability of a failing result due to measurement variability is

$$P\left(X < \text{LTL or UTL} < X\right) = P\left(X < \text{LTL}\right) + P\left(\text{UTL} < X\right) = 1 + \Phi\left(\frac{\text{LTL} - \mu}{\sigma}\right) - \Phi\left(\frac{\text{UTL} - \mu}{\sigma}\right) \quad \textbf{(5.49)}$$

EXAMPLE 5.11

A DC offset measurement is repeated many times, resulting in a series of values having an average of 257 mV. The measurements exhibit a standard deviation of 27 mV. What is the probability that a nonaveraged offset measurement will fail on any given test program execution? Assume an upper test limit of 300 mV and a lower test limit of 200 mV.

Solution:

The probability that the test will lie outside the test limits of 200 and 300 mV is obtained by substituting the test limits into Eq. (5.49).

$$P(X < 200\text{ mV}) + P(300\text{ mV} < X) = 1 + \Phi\left(\frac{200\text{ mV} - 257\text{ mV}}{27\text{ mV}}\right) - \Phi\left(\frac{300\text{ mV} - 257\text{ mV}}{27\text{ mV}}\right)$$

$$= 1 + \Phi(-2.11) - \Phi(1.59)$$

Using Table 5.1, we estimate the cdf values as

$$P(X < 200 \text{ mV}) + P(300 \text{ mV} < X) = 1 + 0.0179 - 0.9452 = 0.0727$$

Here we see that there is a 7.27% chance of failure, even though the true DC offset value is known to be within acceptable limits.

5.9 EFFECTS OF REPRODUCIBILTY AND PROCESS VARIATION ON YIELD

Measured DUT parameters vary for a number of reasons. The factors affecting DUT parameter variation include measurement repeatability, measurement reproducibility, and the stability of the process used to manufacture the DUT. So far we have examined only the effects of measurement repeatability on yield, but the equations in the previous sections describing yield loss due to measurement variability are equally applicable to the total variability of DUT parameters.

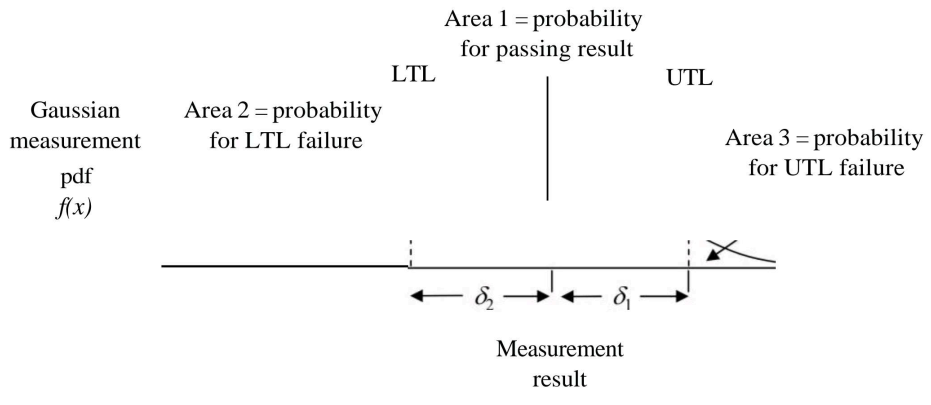
Inaccuracies due to poor tester-to-tester correlation, day-to-day correlation, or DIB-to-DIB correlation appear as reproducibility errors. Reproducibility errors add to the yield loss caused by repeatability errors. To accurately predict yield loss caused by tester inaccuracy, we have to include both repeatability errors and reproducibility errors. If we collect averaged measurements using multiple testers, multiple DIBs, and repeat the measurements over multiple days, we can calculate the mean and standard deviation of the reproducibility errors for each test. We can then combine the standard deviations due to repeatability and reproducibility using the equation

$$\sigma_{tester} = \sqrt{(\sigma_{repeatability})^2 + (\sigma_{reproducibility})^2} \tag{5.50}$$

Yield loss due to total tester variability can then be calculated using the equations from the previous sections, substituting the value of σ_{tester} in place of σ.

The variability of the actual DUT performance from DUT to DUT and from lot to lot also contributes to yield loss. Thus the overall variability can be described using an overall standard deviation, calculated using an equation similar to Eq. (5.50), that is,

Figure 5.18. Probability density with large standard deviation.



$$\sigma_{total} = \sqrt{(\sigma_{repeatability})^2 + (\sigma_{reproducibility})^2 + (\sigma_{process})^2} \tag{5.51}$$

Since σ_{total} ultimately determines our overall production yield, it should be made as small as possible to minimize yield loss. The test engineer must try to minimize the first two standard deviations. The design engineer and process engineer should try to reduce the third.

EXAMPLE 5.12

A six-month yield study finds that the total standard deviation of a particular DC offset measurement is 37 mV across multiple lots, multiple testers, multiple 01B boards, and so on. The standard deviation of the measurement repeatability is found to be 15 mV, while the standard deviation of the reproducibility is found to be 7 mV. What is the standard deviation of the actual OUT-to-OUT offset variability, excluding tester repeatability errors and reproducibility errors? If we could test this device using perfectly accurate, repeatable test equipment, what would be the total yield loss due to this parameter, assuming an average value of 2.430 V and test limits of 2.5 V ± 100 mV?

Solution:

Rearranging Eq. [5.51], we write

$$\sigma_{process} = \sqrt{(\sigma_{total})^2 - (\sigma_{repeatability})^2 - (\sigma_{reproducibility})^2}$$

$$= \sqrt{(37\text{ mV})^2 - (15\text{ mV})^2 - (7\text{ mV})^2}$$

$$= 33\text{ mV}$$

Thus, even if we could test every device with perfect accuracy and no repeatability errors, we would see a OUT-to-OUT variability of $\sigma = 33$ mV. The value of μ is equal to 2.430 V; thus our overall yield loss for this measurement is found by substituting the above values into Eq. [5.49] as

$$P(X < 2.4\text{ V}) + P(2.6\text{ V} < X) = 1 + \Phi\left(\frac{2.4\text{ V} - 2.43\text{ V}}{33\text{ mV}}\right) - \Phi\left(\frac{2.6\text{ V} - 2.43\text{ V}}{33\text{ mV}}\right)$$

$$= 1 + \Phi(-0.91) - \Phi(5.15)$$

From Table 5.1, $\Phi(-0.91) \approx \Phi(-0.9) = 0.1841$, and we estimate $\Phi(5.15) \approx 1$; hence

$$P(X < 2.4\text{ V}) + P(2.6\text{ V} < X) = 1 + 0.1841 - 1 = 0.1841$$

We would therefore expect an 18% yield loss due to this one parameter, due to the fact that the OUT-to-OUT variability is too high to tolerate an average value that is only 30 mV from the lower test limit. Repeatability and reproducibility errors would only worsen the yield loss; so this device would probably not be economically viable. The design or process would have to be modified to achieve an average DC offset value closer to 2.5 V.

The probability that a particular device will pass all tests in a test program is equal to the product of the passing probabilities of each individual test. In other words, if the values P_1, P_2, P_3, ... , Pn represent the probabilities that a particular DUT will pass each of then individual tests in a test program, then the probability that the DUT will pass all tests is equal to

$$P\ (\text{DUT passes all tests}) = P_1 \times P_2 \times P_3 \times \cdots \times Pn \qquad \textbf{(5.52)}$$

Equation (5.52) is of particular significance, because it dictates that each of the individual tests must have a very high yield if the overall production yield is to be high. For example, if each of the 200 tests has a 2% chance of failure, then each test has only a 98% chance of passing. The yield will therefore be $(0.98)^{200}$, or 1.7%! Clearly, a 1.7% yield is completely unacceptable. The problem in this simple example is not that the yield of any one test is low, but that so many tests combined will produce a large amount of yield loss.

EXAMPLE 5.13

A particular test program performs 857 tests, most of which cause little or no yield loss. Five measurements account for most of the yield loss. Using a lot summary and a continue-on-fail test process, the yield loss due to each measurement is found to be:

Test #1: 1%, Test #2: 5%, Test #3: 2.3%, Test #4: 7%, Test #5: 1.5%
All other tests combined 0.5%

What is the overall yield of this lot of material?

Solution:
The probability of passing each test is equal to 1 minus the yield loss produced by that test. The values of P_1, P_2, P_3, ... , P_5 are therefore

$$p1 = 99\%,\ p2 = 95\%,\ p3 = 97.7\%,\ p4 = 93\%,\ p5 = 98.5\%$$

If we consider all other tests to be a sixth test having a yield loss of 0.5%, we get a sixth probability

$$Pb = 99.5\%$$

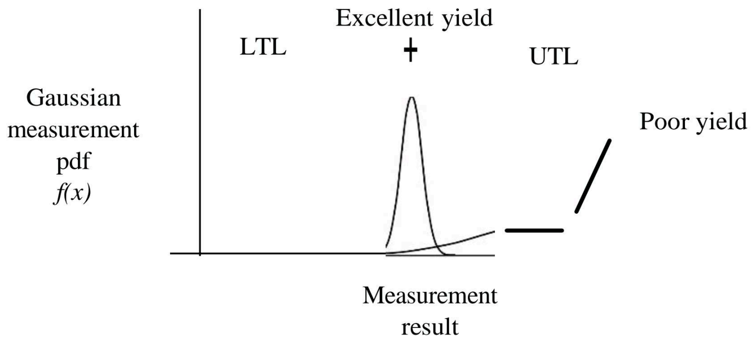
Using Eq. (5.52) we write

$$P(\text{DUT passes all tests}) = 0.99 \times 0.95 \times 0.977 \times 0.93 \times 0.985 \times 0.995 = 0.8375$$

Thus we expect an overall test yield of 83.75%.

Because the yield of each individual test must be very high, a methodology called *statistical process control* (SPC) has been adopted by many companies. The goal of SPC is to minimize the total variability (i.e., to try to make $cr,orat = 0$) and to center the average test result between the upper and lower test limits [i.e. to try to make μ = (UTL+LTL)/2]. Centering and narrowing the measurement distribution leads to higher production yield, since it minimizes the area of the Gaussian pdfs that extend into the failing regions as depicted in Figure 5.19. In the next section, we will briefly

Figure 5.19. OUT-to-OUT mean and standard deviation determine yield.



examine the SPC methodology to see how it can help improve the quality of the manufacturing process, the quality of the test equipment and software, and most important the quality of the devices shipped to the customer.

5.10 STATISTICAL PROCESS CONTROL

5.10.1 Goals of SPC

Statistical process control (SPC) is a structured methodology for continuous process improvement. SPC is a subset of total quality control (TQC), a methodology promoted by the renowned quality expert, Joseph Juran.[7,8] SPC can be applied to the semiconductor manufacturing process to monitor the consistency and quality of integrated circuits.

Exercises

5.26. An AC gain measurement is repeated many times, resulting in a series of values having an average of 2.3 V/V. The measurements exhibit a standard deviation of 0.15 V/V. What is the probability that the gain measurement will fail on any given test program execution? Assume an upper test limit of 2.4 V/V and a lower test limit **of 2.2 V/V.**

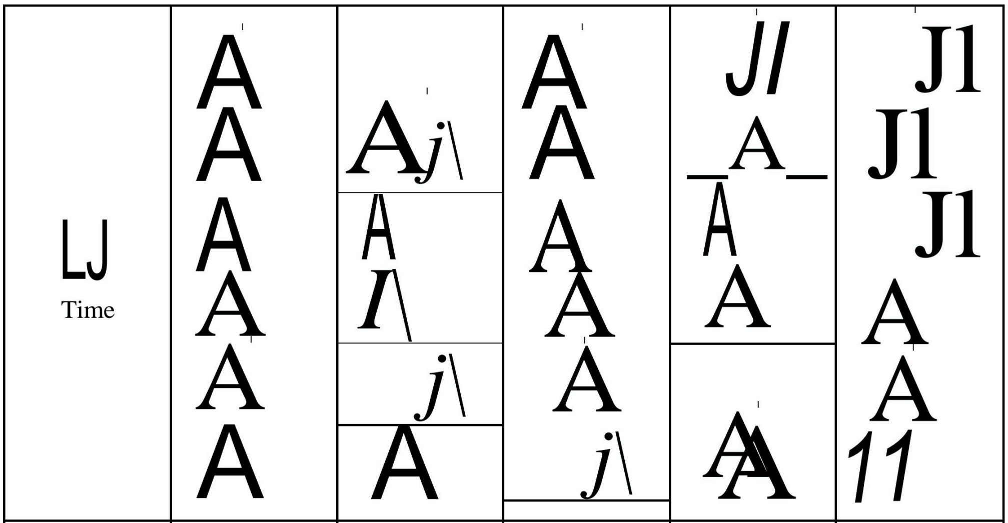
ANS. **0.50.**

5.27. A particular test program performs 600 tests, most of which cause little or no yield loss. Four measurements account for most of the yield loss. The yield loss due to each measurement is found to be: Test #1: 1.5%, Test #2: 4%, Test #3: 5.3%, Test #4: 2%. All other tests combined 5%. What is the overall yield loss of this lot of material?

ANS. Yield loss = 16.63%.

SPC provides a means of identifying device parameters that exhibit excessive variations over time. It does not identify the root cause of the variations, but it tells us when to look for problems. Once an unstable parameter has been identified using SPC, the engineering and manufacturing team searches for the root cause of the instability. Hopefully, the excessive variations can be reduced or eliminated through a design modification or through an improvement in one of the

Figure 5.20. Process stability conclusions.



Centering	Consistent	Inconsistent	Drifting	Consistent	Inconsistent
Variability	Consistent	Consistent	Consistent	Inconsistent	Inconsistent
Conclusion	Stable	Unstable	Unstable	Unstable	Unstable

many manufacturing steps. By improving the stability of each tested parameter, the manufacturing process is brought under control, enhancing the inherent quality of the product.

A higher level of inherent quality leads to higher yields and less demanding test requirements. If we can verify that a parameter almost never fails, then we may be able to stop testing that parameter on a DUT-by-DUT basis. Instead, we can monitor the parameter periodically to verify that its statistical distribution remains tightly packed and centered between the test limits. We also need to verify that the mean and standard deviation of the parameter do not fluctuate wildly from lot to lot as shown in the four rightmost columns of Figure 5.20.* Once the stability of the distributions has been verified, the parameter might only be measured for every tenth device or every hundredth device in production. If the mean and standard deviation of the limited sample set stays within tolerable limits, then we can be confident that the manufacturing process itself is stable. SPC thus allows statistical sampling of highly stable parameters, dramatically reducing testing costs.

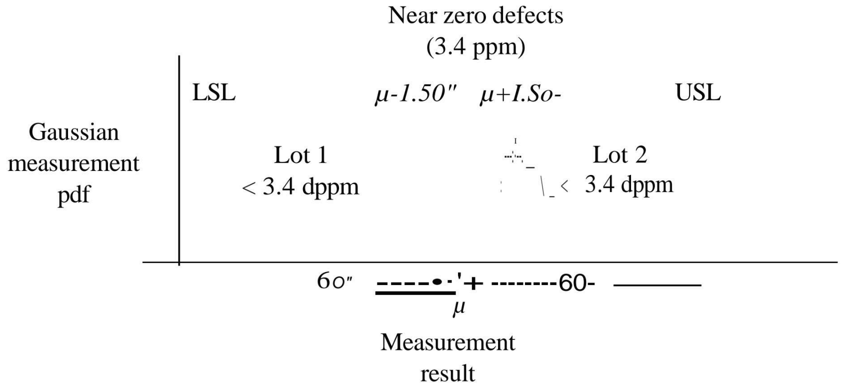
5.10.2 Six-Sigma Quality

If successful, the SPC process results in an extremely small percentage of parametric test failures. The ultimate goal of SPC is to achieve six-sigma quality standards for each specified device parameter. A parameter is said to meet six-sigma quality standards if its standard deviation is no greater than 1/12 of the difference between the upper and lower specification limits *and* the center of its statistical distribution is no more than 1.5σ away from the center of the upper and lower test limits. These criteria are illustrated in Figure 5.21.

---

*The authors acknowledge the efforts of the Texas Instruments SPC Guidelines Steering Team, whose document "Statistical Process Control Guidelines, The Commitment of Texas Instruments to Continuous Improvement Through SPC" served as a guide for several of the diagrams in this section.

Figure 5.169. Six-sigma quality standards lead to low defect rates [< 3.4 defective parts per million).



Six-sigma quality standards result in a failure rate of less than 3.4 defective parts per million (dppm). Therefore, the chance of an untested device failing a six-sigma parameter is extremely low. This is the reason we can often eliminate DUT-by-DUT testing of six-sigma parameters.

5.10.3 Process Capability: C_p, and C_{pk}

Process capability is the inherent variation of the process used to manufacture a product. Process capability is defined as the ±3cr variation of a parameter around its mean value. For example, if a given parameter exhibits a 10-mV standard deviation from DUT to DUT over a period of time, then the process capability for this parameter is defined as 60 mV.

The centering and variation of a parameter are defined using two process stability metrics, C_p and C_{pk}. The process potential index, C_p is the ratio between the range of passing values and the process capability

$$\frac{USL-LSL}{6cr} \tag{5.53}$$

C_p indicates how tightly the statistical distribution of measurements is packed, relative to the range of passing values. A very large C_p value indicates a process that is stable enough to give high yield and high quality, while a C_p less than 2 indicates a process stability problem. It is impossible to achieve six-sigma quality with a C_p less than 2, even if the parameter is perfectly centered. For this reason, six-sigma quality standards dictate that all measured parameters must maintain a C_p of 2 or greater in production.

The process capability index, C_{pk}' measures the process capability with respect to centering between specification limits

$$C_{pk} = C_p (1-k) \tag{5.54}$$

where

$$k = \frac{|T-\mu|}{0.5(USL - LSL)} \tag{5.55}$$

Here Tis the specification target (ideal measured value) andμ is the average measured value. The target value Tis generally placed in the middle of the specification limits, defined as

$$T = \frac{USL+LSL}{2} \qquad (5.56)$$

For one-sided specifications, such as a signal-to-distortion ratio test, we only have an upper or lower specification limit. Therefore, we have to use slightly different calculations for C_p and C_pk. In the case of only the upper specification limit being defined, we use

$$C_{pk} = C_p = \frac{USL-\mu}{30"} \qquad (5.57)$$

Alternatively, with only the lower specification limit defined, we use

$$C_{pk} = C_p = \frac{\mu-LSL}{30"} \qquad (5.58)$$

The value of *Cpk* must be 1.5 or greater to achieve six-sigma quality standards as shown in Figure 5.21.

EXAMPLE 5.14

The values of an AC gain measurement are collected from a large sample of the DUTs in a production lot. The average reading is 0.991 V/V and the upper and lower specification limits are 1.050 and 0.950 V/V, respectively. The standard deviation is found to be 0.0023 V/V. What is the process capability and the values of C_p and C_pk for this lot? Does this lot meet six-sigma quality standards?

Solution:

The process capability is equal to 6 sigma, or 0.0138 V/V. The values of C_p and C_pk are given by Eqs. (5.53]-(5.55]:

$$C_p = \frac{USL-LSL}{6<:5} = \frac{1.050-0.950}{0.0138} = 7{-}245$$

$$k = \frac{|T-\mu|}{0.5(USL-LSL)} \quad \frac{11-0.9911}{0.5(1.050-0.950)} = 0\ 18\ .$$

$$C_{pk} = C_P(1-k) = 5.94$$

This parameter meets six-sigma quality requirements, since the values of C_p is greater than 2 and C_{pk} is greater than 1.5.

5.10.4 Gauge Repeatability and Reproducibility

As mentioned previously in this chapter, a measured parameter's variation is partially due to variations in the materials and the process used to fabricate the device and partially due to the tester's repeatability errors and reproducibility errors. In the language of SPC, the tester is known as a

Table 5.3. %GRR Acceptance Criteria

Measurement Cp	%GRR	Rating
1	100	Unacceptable
3	33	Unacceptable
5	20	Marginal
10	10	Acceptable
50	2	Good
100	1	Excellent

gauge. Before we can apply SPC to a manufacturing process, we first need to verify the accuracy, repeatability, and reproducibility of the gauge. Once the quality of the testing process has been established, the test data collected during production can be continuously monitored to verify a stable manufacturing process.

Gauge repeatability and reproducibility, denoted GRR, is evaluated using a metric called *measurement* C_p. We collect repeatability data from a single DUT using multiple testers and different DIBs over a period of days or weeks. The composite sample set represents the combination of tester repeatabilty errors and reproducibility errors [as described by Eq. (5.50)]. Using the composite mean and standard deviation, we calculate the measurement C_p using Eq. (5.53). The gauge repeatability and reproducibility percentage (precision-to-tolerance ratio) is defined as

$$\%GRR = \frac{100}{\text{measurement } C_P} \tag{5.59}$$

The general criteria for acceptance of gauge repeatability and reproducibility are listed in Table 5.3.

PROBLEMS

5.1. If 20,000 devices are tested with a yield of 98%, how many devices failed the test?

5.2. A new product was launched with 20,000 sales over a one-year time span. During this time, three devices were returned to the manufacturer even though an extensive test screening procedure was in place. What is the defect level associated with this testing procedure in parts per million?

5.3. A 55-mV signal is measured with a meter 10 times, resulting in the following sequence of readings: 57 mV, 60 mV, 49 mV, 58 mV, 54 mV, 57 mV, 55 mV, 57 mV, 48 mV, 61 mV. What is the average measured value? What is the systematic error?

5.4. A DC voltmeter is rated at 14 bits of resolution and has a full-scale input range of ±5 V. Assuming the meter's ADC is ideal, what is the maximum quantization error that we can expect from the meter? What is the error as a percentage of the meter's full-scale range?

5.5. A 100-mV signal is to measured with a worst-case error of ±10 µV. A DC voltmeter is set to a full-scale range of ±1 V. Assuming that quantization error is the only source of inaccuracy in this meter, how many bits of resolution would this meter need to have to make the required measurement? If the meter in our tester only has 14 bits of resolution but has binary-weighted range settings (i.e., ±1 V, ±500 mV, ±250 mV, etc.), how would we make this measurement?

5.6. A voltmeter is specified to have an accuracy error of ±0.1% of full-scale range on a ±1-V scale. If the meter produces a reading of 0.323 V DC, what is the minimum and maximum DC levels that might have been present at the meter's input during this measurement?

5.7. A series of 100 gain measurements was made on a DUT whereby the distribution was found to be Gaussian with mean value of 10.1 V/V and a standard deviation of 0.006 V/V. Write an expression for the pdf of these measurements.

5.8. A series of 100 gain measurements was made on a DUT whereby the distribution was found to be Gaussian with mean value of 10.1 V/V and a standard deviation of 0.006 V/V. If this experiment is repeated, write an expression for the pdf of the mean values of each of these experiments?

5.9. A series of 100 gain measurements was made on a DUT whereby the distribution was found to be Gaussian with mean value of 10.1 VN and a standard deviation of 0.006 V/V. If this experiment is repeated and the mean value is compared to a reference gain value of 10 V*N,* what is the mean and standard deviation of the error distribution that results? Write an expression for the pdf of these errors.

5.10. A series of 100 gain measurements was made on a DUT whereby the distribution was found to be Gaussian with mean value of 10.1 V/V and a standard deviation of 0.006 V/V. If this experiment and the mean value is compared to a reference value of 10 V*N,* in what range will the expected value of the error lie for a 99.7% conference interval.

5.11. A meter reads -1.039 V and 1.121 V when connected to two highly accurate reference levels of -1 V and 1 V, respectively. What is the offset and gain of this meter? Write the calibration equation for this meter.

5.12. A DC source is assumed characterized by a third-order equation of the form: $V_{souRcED} = 0.004 + V_{PROGRAMMED} + 0.00lV_{P\,ocRAMMED} - 0.007V_{\}RocRAMMED}$ and is required to generate a DC level of 1.25 V. However, when programmed to produce this level, 1.242 V is measured. Using iteration, determine a value of the programmed source voltage that will establish a measured voltage of 1.25 V to within a ± 0.5 mV accuracy.

5.13. AnAWG has a gain response described by $G(f)=\frac{1}{\sqrt{1+\left(\frac{f}{4000}\right)^{2}}}$ and is to generate three tones at frequencies of 1, 2, and 3 kHz. What are the gain calibration factors? What voltage levels would we request if we wanted an output level of 500 mV RMS at each frequency?

5.14. Several DC measurements are made on a signal path that contains a filter and a buffer amplifier. At input levels of 1 and 3 V, the output was found to be 1.02 and 3.33 V, respectively. Assuming linear behavior, what is the gain and offset of this filter-buffer stage?

5.15. Using the setup and results of Problem 5.14, what is the calibrated level when a 2.13 V level is measured at the filter-buffer output? What is the size of the uncalibrated error?

5.16. A simple RC low-pass circuit is constructed using a 1-kQ resistor and a 10-μF capacitor. This RC circuit is used to filter the output of a DDT containing a noisy DC signal. If the DUT's noise voltage has a constant spectral density of 100 nV/,Jtt;,, what is the RMS noise voltage that appears at the output of the RC filter? If the we decrease the capacitor value to 2.2 μF, what is the RMS noise voltage at the RC filter output?

5.17. Assume that we want to allow the RC filter in Problem 5.16 to settle to within 0.2% of its final value before making a DC measurement. How much settling time does the first RC filter in Problem 5.16 require? Is the settling time of the second RC filter greater or less than that of the first filter?

5.18. A DC meter collects a series of repeated offset measurements at the output of a DDT. A first-order low-pass filter such as the first one described in Problem 5.16 is connected between the DDT output and the meter input. A histogram is produced from the repeated measurements. The histogram shows a Gaussian distribution with a 50-mV difference between the maximum value and minimum value. It can be shown that the standard deviation, CT, of the histogram of a repeated series of identical DC measurements on one DDT is proportional to the RMS noise at the meter's input. Assume that the difference between the maximum and minimum measured values is roughly equal to *6a*. How much would we need to reduce the cutoff frequency of the low-pass filter to reduce the nonrepeatability of the measurements from 50 to 10 mV? What would this do to our test time, assuming that the test time is dominated by the settling time of the low-pass filter?

5.19. The DDT in Problem 5.16 can be sold for $1.25, assuming that it passes all tests. If it does not pass all tests, it cannot be sold at all. Assume that the more repeatable DC offset measurement in Problem 5.16 results in a narrower guardband requirement, causing the production yield to rise from 92% to 98%. Also assume that the cost of testing is known to be 3.5 cents per second and that the more repeatable measurement adds 250 ms to the test time. Does the extra yield obtained with the lower filter cutoff frequency justify the extra cost of testing resulting from the filter's longer settling time?

5.20. A series of DC offset measurements reveal an average value of 10 mV and a standard deviation of 11 mV. If our specification limits were 0 ± 50 mV, where would we have to set our 3a guardbanded upper and lower test limits? If *6a* guardbands are desired, how many times would we have to average the measurement to achieve guardbands of 20 mV?

5.21. A DC offset measurement is repeated many times, resulting in a series of values having an average of -100 mV. The measurements exhibit a standard deviation of 38 mV. What is the probability that the offset measurement will fail on any given test program execution? Assume an upper test limit of 0 mV and a lower test limit of -150 mV. Provide a sketch of the pdf, label critical points, and highlight the area under the pdf that corresponds to the probability of interest.

5.22. A gain measurement is repeated many times, resulting in a series of values having an average of 6.5 V/V. The measurements exhibit a standard deviation of 0.05 VN. If our specification limits are 6.0 ± 0.5 V/V, where would we have to set our 3 *a* guardbanded upper and lower test limits? If *6a* guardbands are desired, how many times would we have to average the measurement to achieve guardbands of 0.1 V/V?

5.23. A DC offset test is performed on a DUT with lower and upper specification limits of -12 mV and 12 mV, respectively. The expected RMS level of the noise present during the test is 1.5 mV. If a defect level of less than 200 ppm is required, what should be the test limits?

5.24. A device is expected to exhibit a worst-case offset voltage of ±10 mV and is to be measured using a voltmeter having an accuracy of only ±500 μV. Where should the guardbanded test limits be set?

5.25. The guardband of a particular measurement is 0.2 V/V and the test limits are set to 6.1 VN and 6.2 V/V. What are the original device specification limits?

5.26. A series of DC measurements reveal the following list of values:

{0 mV, -10 mV, 1.5 mV, 9.5 mV, -8.5 mV, 13.2 mV,
18.5 mV, -17.2 mV, 5.3 mV, and 6.2 mV}

If our specification limits were 0 ± 50 mV, where would we have to set our 3 *a* guardband upper and lower test limits? Provide a sketch to illustrate the probability density function and show the test limits. If *6a* guardbands are desired, how many times would we have to average the measurement to achieve guardbands of 24 mV?

5.27. The following contains a list of output voltage values from a DUT together with their actual measured values (i.e., sets of (true value, measured value)):

{(1.9, 1.81), (2.1, 1.75), (2.1,1.77), (1.8, 1.79), (1.9, 1.71), (2.1, 1.95),
(2.2, 2.11), (1.7,1.89), (1.5, 1.7)}

If the upper specification limit is 2 V and the guardbanded upper test limit is set to 1.8 V, answer the following questions:

(a) How many good devices are rejected on account of measurement error?
(b) How many devices escape the test?
(c) If the upper test limit is reduced to 1.74 V, how many devices escape on account of measurement error?

5.28. An AC gain measurement is repeated many times, resulting in a series of values having an average of 0.99 V/V. The measurements exhibit a standard deviation of 0.2 VN. What is the probability that the gain measurement will fail on any given test program execution? Assume an upper test limit of 1.2 V/V and a lower test limit of 0.98 V/V. Provide a sketch of the pdf, label critical points, and highlight the area under the pdf that corresponds to the probability of interest.

5.29. The standard deviation of a measurement repeatability is found to be 12 mV, while the standard deviation of the reproducibility is found to be 8 mV. Determine the standard deviation of the tester's variability. If process variation contributes an additional 10 mV of uncertainity to the measurement, what is the total standard deviation of the overall measurement?

5.30. An extensive study of yield finds that the total standard deviation of a particular DC offset measurement is 25 m V across multiple lots, multiple testers, multiple DIB boards, and so on. The standard deviation of the measurement repeatability is found to be 19 m V, while the standard deviation of the reproducibility is found to be 11 mV. What is the standard deviation of the actual DUT-to-DUT offset variability, excluding tester repeatability errors and reproducibility errors? If we could test this device using perfectly accurate, repeatable