Flowing Expectations for Minimum Probability of Conformance between Engineering and Operations

Dan Fitzsimmons[1]

[1]Dan K. Fitzsimmons, Boeing Commercial Airplanes, P.O.Box 3707, MC 08-51,
Seattle, WA 98124; dan.k.fitzsimmons@boeing.com

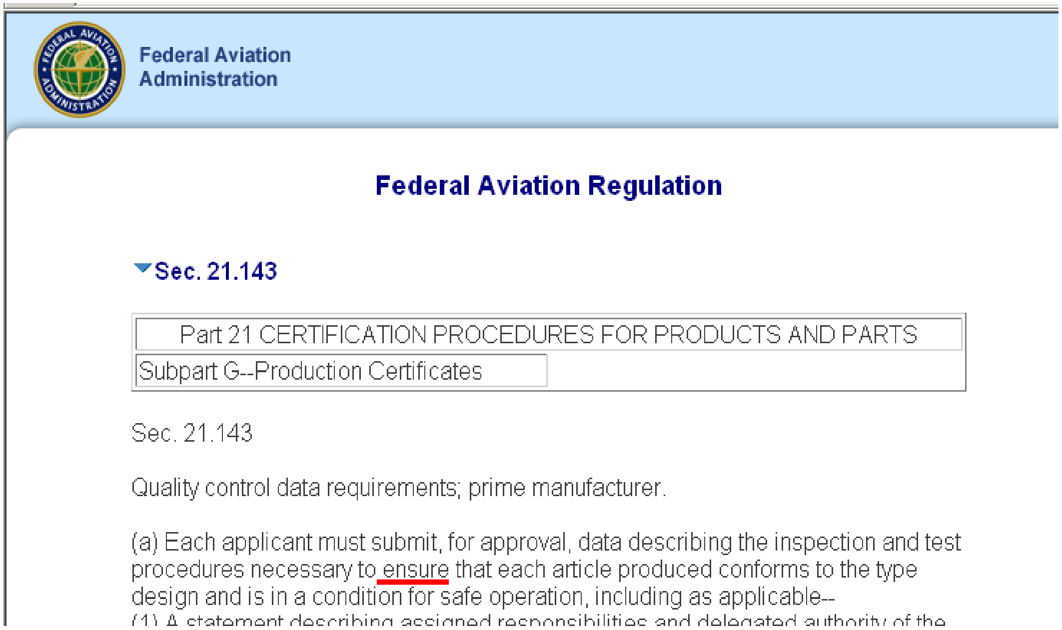
Abstract

- Aerospace products, especially commercial aircraft, require exceptionally high reliability.
    - US Federal Aviation Administration Advisory Circular recommendations for design failure probabilities are in many cases less than 10^ -9.
- Requirements this tight call for achieving high probability of conformance values for the product stream going into those aerospace products.
- Probability of conformance values in the design must be achieved by the operations people building the product.
- This makes it necessary to communicate what those operations have historically achieved back to the design community.

This paper presents methods that have been used in the aerospace industry to accomplish this communication, without either set of participants being trained statisticians.

Key Words: AQL, AOQL, Expected Outgoing Quality, Initial Reliability Requirement, minimum probability of conformance, Statistical Engineering

1. FAA Requirements



Federal Aviation Regulation

Sec. 21.143

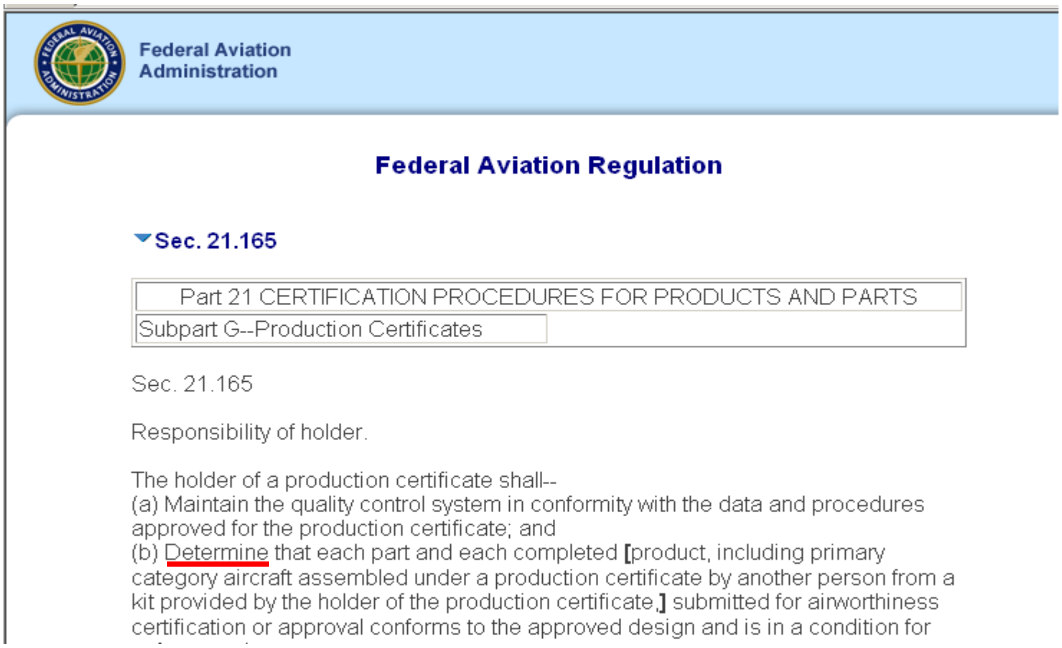
Part 21 CERTIFICATION PROCEDURES FOR PRODUCTS AND PARTS
Subpart G--Production Certificates

Sec. 21.143

Quality control data requirements; prime manufacturer.

(a) Each applicant must submit, for approval, data describing the inspection and test procedures necessary to ensure that each article produced conforms to the type design and is in a condition for safe operation, including as applicable--
(1) A statement describing assigned responsibilities and delegated authority of the

Figure 1: Excerpt from Federal Aviation Regulation 21.143



Federal Aviation Regulation

Sec. 21.165

Part 21 CERTIFICATION PROCEDURES FOR PRODUCTS AND PARTS

Subpart G--Production Certificates

Sec. 21.165

Responsibility of holder.

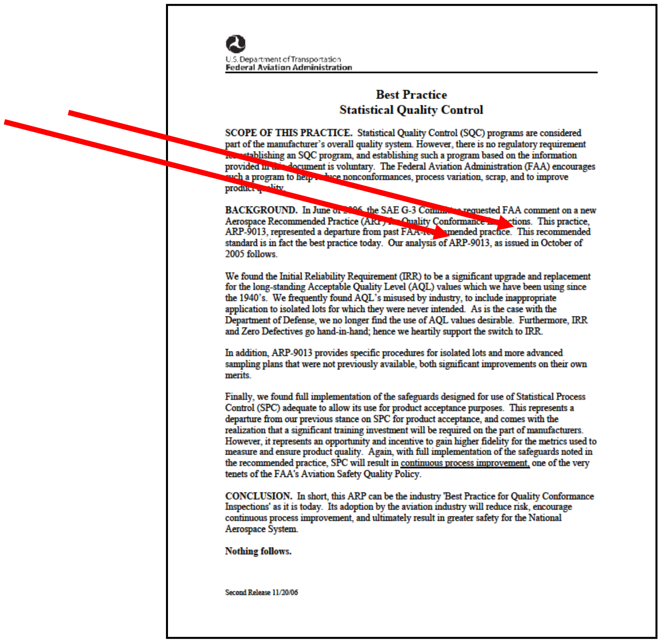
The holder of a production certificate shall--
(a) Maintain the quality control system in conformity with the data and procedures approved for the production certificate; and
(b) Determine that each part and each completed [product, including primary category aircraft assembled under a production certificate by another person from a kit provided by the holder of the production certificate,] submitted for airworthiness certification or approval conforms to the approved design and is in a condition for

Figure 2: Excerpt from Federal Aviation Regulation 21.165

So, PRECISELY, what do “ensure” and “determine” mean? The FAA has avoided stating explicitly a precise technical definition to these words. However, they have given statements that logically lead to useful conclusions.

1.1 FAA “Best Practice in Statistical Quality Control”, 2006-2011

This “Best Practice” was first published in 2001 as guidance to their auditors on statistical methods that met the FAR. In 2006, it was reduced to this one-page statement designating SAE ARP9013 as the “Best Practice for Quality Conformance Inspections”.



U.S. Department of Transportation
Federal Aviation Administration

Best Practice
Statistical Quality Control

SCOPE OF THIS PRACTICE. Statistical Quality Control (SQC) programs are considered part of the manufacturer’s overall quality system. However, there is no regulatory requirement for establishing an SQC program, and establishing such a program based on the information provided in this document is voluntary. The Federal Aviation Administration (FAA) encourages such a program to help reduce nonconformances, process variation, scrap, and to improve product quality.

BACKGROUND. In June of 2006, the SAE G-3 Committee requested FAA comment on a new Aerospace Recommended Practice (ARP) for Quality Conformance Inspections. This practice, ARP-9013, represented a departure from past FAA-recommended practice. This recommended standard is in fact the best practice today. Our analysis of ARP-9013, as issued in October of 2005 follows.

We found the Initial Reliability Requirement (IRR) to be a significant upgrade and replacement for the long-standing Acceptable Quality Level (AQL) values which we have been using since the 1940’s. We frequently found AQL’s misused by industry, to include inappropriate application to isolated lots for which they were never intended. As is the case with the Department of Defense, we no longer find the use of AQL values desirable. Furthermore, IRR and Zero Defectives go hand-in-hand; hence we heartily support the switch to IRR.

In addition, ARP-9013 provides specific procedures for isolated lots and more advanced sampling plans that were not previously available, both significant improvements on their own merits.

Finally, we found full implementation of the safeguards designed for use of Statistical Process Control (SPC) adequate to allow its use for product acceptance purposes. This represents a departure from our previous stance on SPC for product acceptance, and comes with the realization that a significant training investment will be required on the part of manufacturers. However, it represents an opportunity and incentive to gain higher fidelity for the metrics used to measure and ensure product quality. Again, with full implementation of the safeguards noted in the recommended practice, SPC will result in continuous process improvement, one of the very tenets of the FAA's Aviation Safety Quality Policy.

CONCLUSION. In short, this ARP can be the industry 'Best Practice for Quality Conformance Inspections' as it is today. Its adoption by the aviation industry will reduce risk, encourage continuous process improvement, and ultimately result in greater safety for the National Aerospace System.

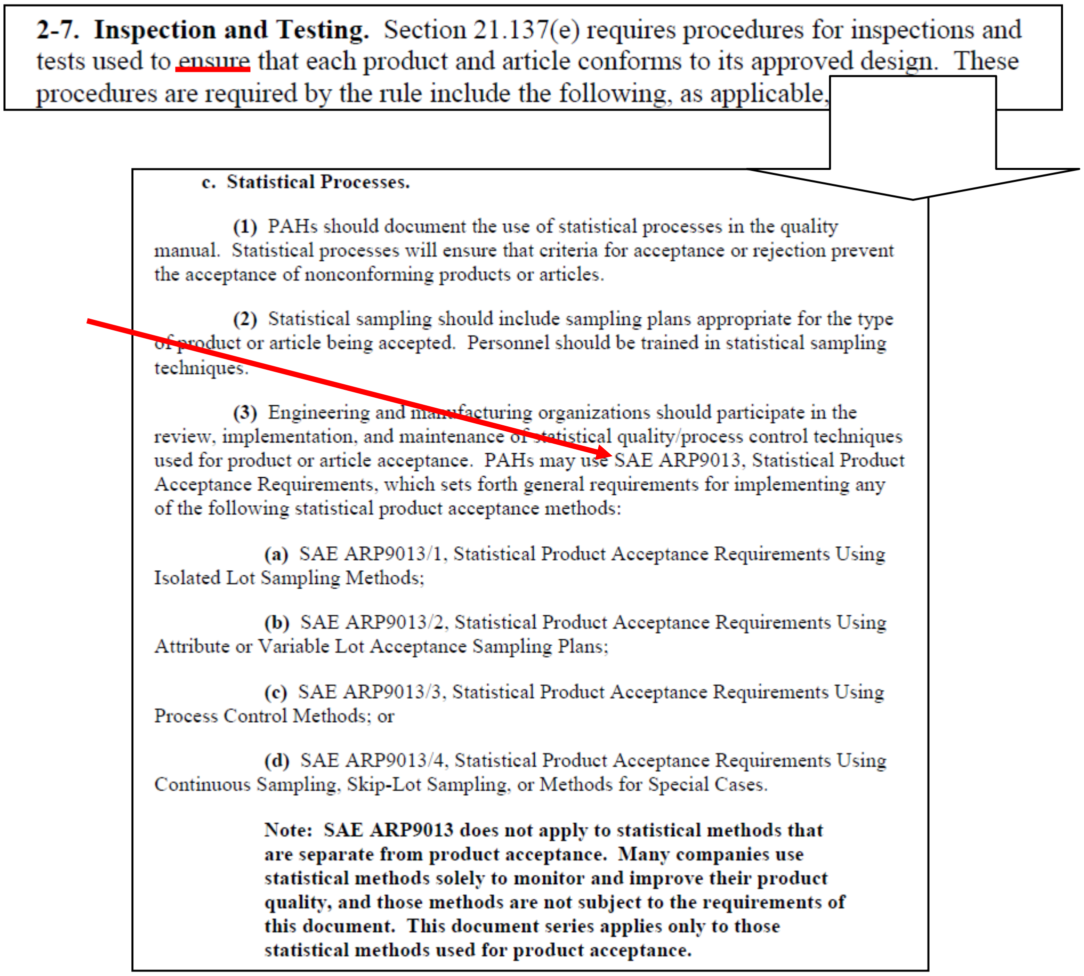
Nothing follows.

Second Release 11/20/06

Figure 3: FAA “Best Practice” / “Statistical Quality Control”, 2006

In 2011, the FAA discontinued using “Best Practice” statements to guide their auditors. This one was replaced by AC 21-43:

1.2 FAA Advisory Circular 21-43 (2011-current)



2-7. Inspection and Testing. Section 21.137(e) requires procedures for inspections and tests used to ensure that each product and article conforms to its approved design. These procedures are required by the rule include the following, as applicable.

c. Statistical Processes.

(1) PAHs should document the use of statistical processes in the quality manual. Statistical processes will ensure that criteria for acceptance or rejection prevent the acceptance of nonconforming products or articles.

(2) Statistical sampling should include sampling plans appropriate for the type of product or article being accepted. Personnel should be trained in statistical sampling techniques.

(3) Engineering and manufacturing organizations should participate in the review, implementation, and maintenance of statistical quality/process control techniques used for product or article acceptance. PAHs may use SAE ARP9013, Statistical Product Acceptance Requirements, which sets forth general requirements for implementing any of the following statistical product acceptance methods:

(a) SAE ARP9013/1, Statistical Product Acceptance Requirements Using Isolated Lot Sampling Methods;

(b) SAE ARP9013/2, Statistical Product Acceptance Requirements Using Attribute or Variable Lot Acceptance Sampling Plans;

(c) SAE ARP9013/3, Statistical Product Acceptance Requirements Using Process Control Methods; or

(d) SAE ARP9013/4, Statistical Product Acceptance Requirements Using Continuous Sampling, Skip-Lot Sampling, or Methods for Special Cases.

Note: SAE ARP9013 does not apply to statistical methods that are separate from product acceptance. Many companies use statistical methods solely to monitor and improve their product quality, and those methods are not subject to the requirements of this document. This document series applies only to those statistical methods used for product acceptance.

Figure 4: Excerpts from AC 21-43

1.3 What does ARP9013 say about “ensure” and “determine”?

In its definition for “Quality Parameter”, ARP9013 says that satisfying a Quality Parameter provides a basis to say that the conformance of the product has been “ensured” & “determined”.

ARP9013 explains in the Requirements section that a Quality Parameter is a statistic that protects the consumer by delivering the necessary probability of conformance. ARP9013 includes a Figure B1 illustrating “typical” probability of conformance values used in the legacy of aerospace. These turn out to be significant, so let’s look at their history in more detail.

2. What came before ARP9013?

2.1 The AQL

From 1942 to 1996, aerospace industry worldwide and both for commercial and military products nearly universally used US mil-specs. Nearly every mil-spec designated one or more AQL values for its quality requirements. The "AQL" defined acceptable fractions defective, and thereby directly impacted reliability engineering as well.

Usually the AQL values were between 0.65% and 6.5%. The most common values were 1%, 1.5%, 2.5%, and 4%. These numbers actually reflected protection for the producer of a product rather than the consumer – the military sampling plans were designed to have a high probability of ACCEPTING those defect rates.

However, the need was to protect the consumer…

If the military specifications for statistical product acceptance were followed accurately, the military wrote that these AQL values produced AOQL values that were "close": 1% AQL = 1.22% AOQL, 4% AQL = 4.94% AOQL, etc.

The AOQL was a long-term control of consumer risk. It was the worst average outgoing quality that a sampling plan would allow to pass by a perfect inspector over a "long" production run (i.e., asymptotically). Over the long term, a minimum probability of conformance (MPC) could be met by having the AOQL be less than its complement (that is, AOQL <= 1-MPC).

2.2 The Usefulness of the AQLs

The AQL value in a specification was widely used as the index probability of non-conformance.

Responding to that –

- Design engineers who were considering calling that specification out on their drawings could evaluate whether they had enough margins, redundancies, etc., so that users would be safe with that AQL quality.
- Manufacturing engineers could consider the AQL in choosing appropriate tooling and production methods to achieve that probability of conformance.
- Manufacturing management could consider the AQL in choosing skill levels and appropriate levels of training for production employees.
- Quality engineers could consider the AQL in choosing corresponding systems for verification, including not only sampling risk, but also inspection tooling, measurement accuracy, calibration frequencies, etc.
- Quality management could consider the AQL in choosing skill levels and appropriate training, certification, auditing, and delegation alternatives.

2.3 The Bad Thing About AQLs:

The AQL was not a max probability of non-conformance.

When the long-term average quality was the only concern, and the military sampling standards were used correctly, the AQL was ~20% less than the max probability of nonconformance. However, controlling a long-term average does not prevent local clusters of nonconformances. For safety applications, preventing large local clusters of escapes can be more important than controlling the long term average nonconformance rate.

And, OEMs with qualified statistical quality engineers reported the correct use of AQL-based sampling standards was difficult:

> One statistical quality engineer I spoke with during the legacy period reported from careful records over a period of several years that 95% of the sampling plans submitted by that company's suppliers failed to correctly use the AQL-based sampling plans. Among more than a dozen qualified reviewers I have informally surveyed since then, all but one have reported similar experience.

To cap it off, the textbook definition of the AQL was an index of "producer's risk" as opposed to "consumer's risk".

2.4 US Gov't. Statements About the "AQL"

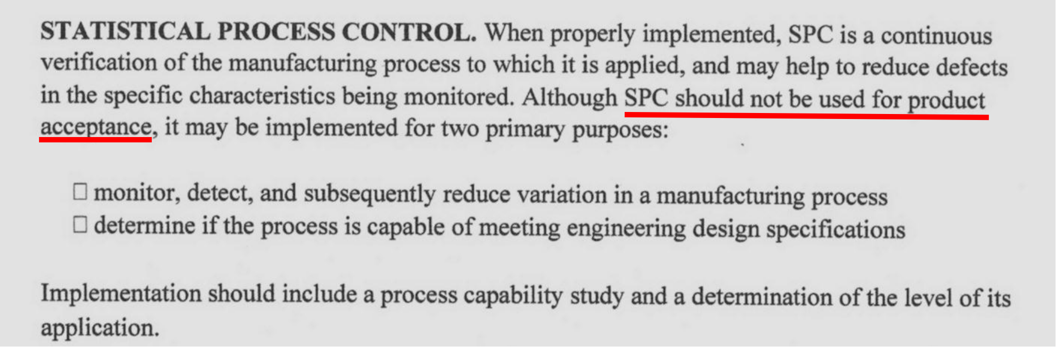
> MIL-HDBK-1916 (1999): "The DoD has stated that Military and Federal Specifications that prescribe fixed levels of nonconformances, such as Acceptable Quality Levels (AQLs) and Lot Tolerance Percent Defectives (LTPDs), inhibit quality improvements and effective competition based on excellence, and should be eliminated."
>
> FAA "Best Practice" 2006 -- "We found the Initial Reliability Requirement (IRR) to be a significant upgrade and replacement for the long-standing Acceptable Quality Level (AQL) values which we have been using since the 1940's. We frequently found AQL's misused by industry, to include inappropriate application to isolated lots for which they were never intended. As is the case with the Department of Defense, we no longer find the use of AQL values desirable. Furthermore, IRR and Zero Defectives go hand-in-hand; hence we heartily support the switch to IRR. "

2.5 What alternatives did Quality Engineering have?

- <u>LTPD</u>? – Also known as LQ – Specifically prohibited in the same DoD letters as the AQL
- <u>RQL</u>? – Tarred with the same brush; also, increased sample sizes by 5x – 10x.
- <u>AOQL</u>? – A better choice, even though still a "specification requirement for a fixed level of defects." However, the AOQL was a long term average for a process, insensitive to short surges in defect rates. And it required rectification, so it was unusable for destructive test, process monitoring, receiving inspection applications.
- <u>Cpk</u>? – This was what the DoD initially said they encouraged …

2.6 Was the Cpk Realistic?

Industry advised the DoD against replacing all their tens of thousands of AQL callouts with blanket Cpk requirements. Just one typical machine shop could be forced to maintain over 1 million SPC charts. Plus, most of the aerospace industry still needed to satisfy the FAA, and they said (in the then-current "Best Practice Statistical Quality Control"):



STATISTICAL PROCESS CONTROL. When properly implemented, SPC is a continuous verification of the manufacturing process to which it is applied, and may help to reduce defects in the specific characteristics being monitored. Although SPC should not be used for product acceptance, it may be implemented for two primary purposes:

☐ monitor, detect, and subsequently reduce variation in a manufacturing process
☐ determine if the process is capable of meeting engineering design specifications

Implementation should include a process capability study and a determination of the level of its application.

Figure 5: FAA “Best Practice” / “Statistical Quality Control”, 2001

3. What About Cost-Optimal Inspection?

There is a way to calculate the cost-optimal amount of inspection. For applications where safety is not an issue, Mood’s Theorem calculates the cost-optimal quality parameter from:

- Marginal cost per inspection, K_i
- Marginal cost per escape, K_e

Both values can usually be estimated quickly to the nearest order of magnitude, but not more precisely than that. The IRR (or MPC) should then be $1\text{-}K_i/K_e$ (see W.Edwards Deming’s “Out of the Crisis” chapter 15 for more details and illustrations).

However, in commercial aerospace, many parts support safety. So …

4. Convergence with Reliability

4.1 The Need

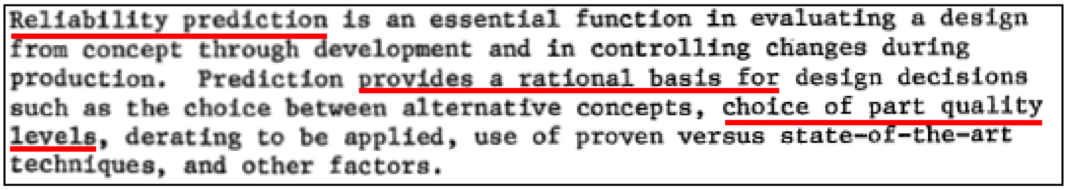
The various functions within the aerospace industry needed that common value represented by the AQL. Design engineers, manufacturing engineers, training functions, quality engineers had choices to make that needed to integrate together.

However, each function actually needed a maximum probability of non-conformance, or equivalently, a minimum probability of conformance.

Enter the discussion with the discipline of Reliability Engineering:

4.2 Legacy Reliability Standards

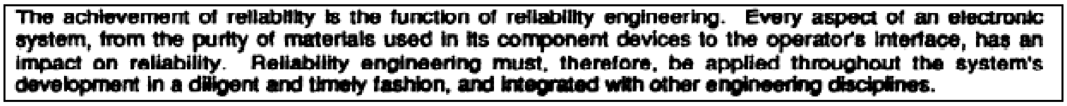
MIL-STD-756 “RELIABILITY MODELING AND PREDICTION (1961-1998):



Reliability prediction is an essential function in evaluating a design from concept through development and in controlling changes during production. Prediction provides a rational basis for design decisions such as the choice between alternative concepts, choice of part quality levels, derating to be applied, use of proven versus state-of-the-art techniques, and other factors.

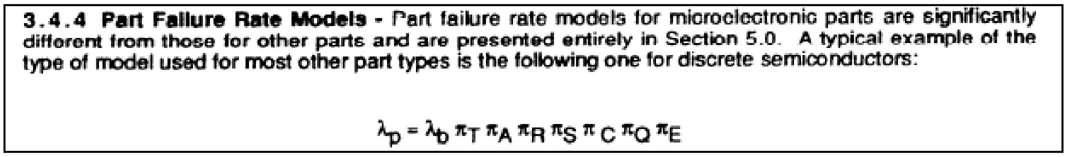
Figure 6: Excerpt from MIL-STD-756 to show the Reliability Engineering viewpoint

MIL-HDBK-217 “RELIABILITY PREDICTION OF ELECTRONIC EQUIPMENT” (1962-CURRENT):



The achievement of reliability is the function of reliability engineering. Every aspect of an electronic system, from the purity of materials used in its component devices to the operator's interface, has an impact on reliability. Reliability engineering must, therefore, be applied throughout the system's development in a diligent and timely fashion, and integrated with other engineering disciplines.

...



3.4.4 Part Failure Rate Models - Part failure rate models for microelectronic parts are significantly different from those for other parts and are presented entirely in Section 5.0. A typical example of the type of model used for most other part types is the following one for discrete semiconductors:

$$\lambda_p = \lambda_b \pi_T \pi_A \pi_R \pi_S \pi_C \pi_Q \pi_E$$

Figure 7: Excerpt from MIL-STD-217, again to show the Reliability viewpoint

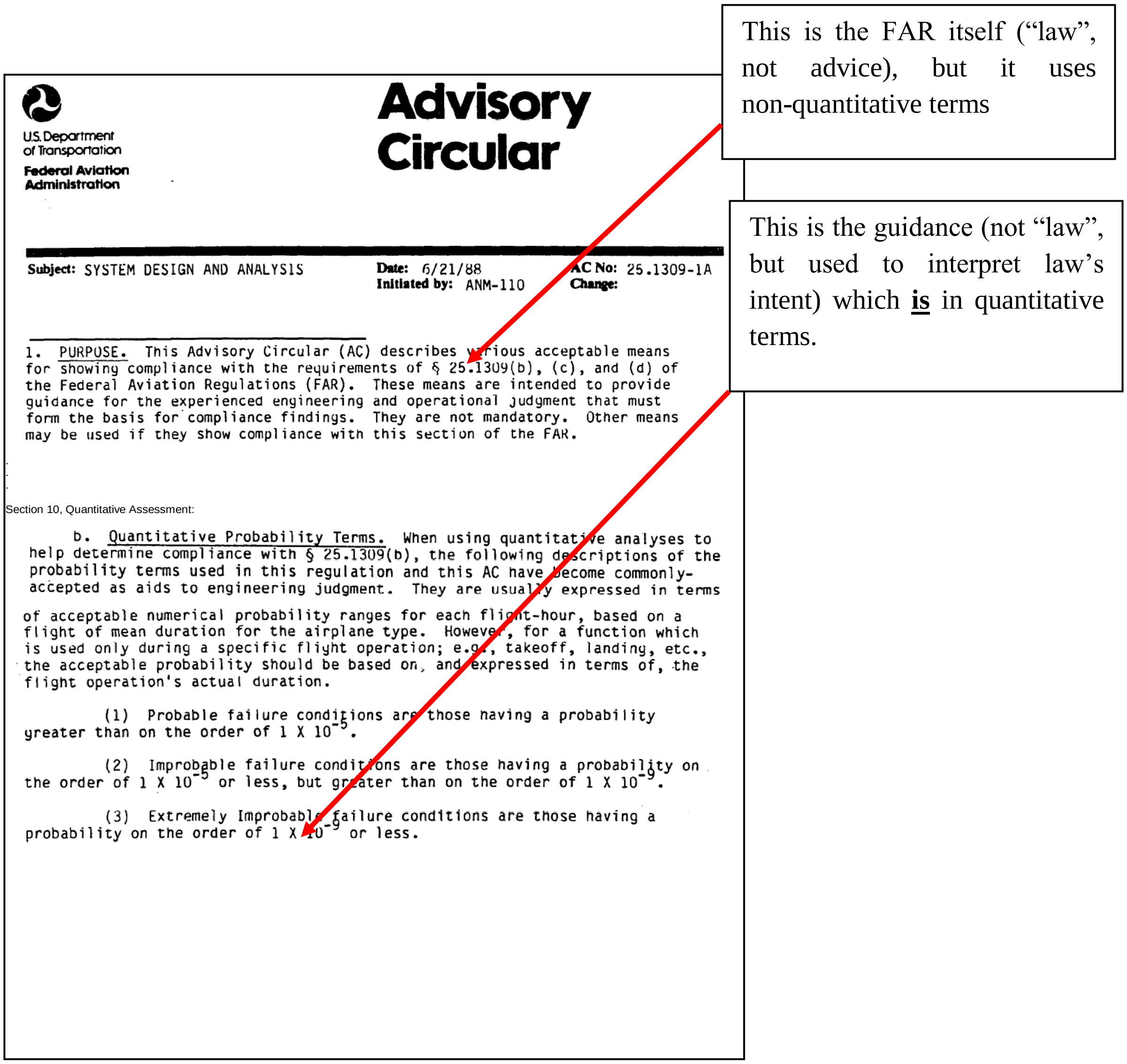
where λ_p is the part failure rate, λ_b is the base failure rate, and the π_ terms modify the base rate to reflect the results of various environmental and other factors. Π_Q in particular adjusts the reliability prediction for part quality levels, and correspondingly, requires compensation in other design elements if quality is low.

4.3 So how does Statistical Engineering intend for QA respond to that?

When safety may be an issue, QA should consult with the Type Design owners. This often brings in Reliability Engineering, and with them, both:

- The maximum failure rate allowed for the relevant assembly, system or structure
- The fleet experience data for the relevant assembly, system or structure.

How important is this? …



U.S. Department of Transportation
Federal Aviation Administration

Advisory Circular

Subject: SYSTEM DESIGN AND ANALYSIS
Date: 6/21/88
Initiated by: ANM-110
AC No: 25.1309-1A
Change:

1. PURPOSE. This Advisory Circular (AC) describes various acceptable means for showing compliance with the requirements of § 25.1309(b), (c), and (d) of the Federal Aviation Regulations (FAR). These means are intended to provide guidance for the experienced engineering and operational judgment that must form the basis for compliance findings. They are not mandatory. Other means may be used if they show compliance with this section of the FAR.

Section 10, Quantitative Assessment:

b. Quantitative Probability Terms. When using quantitative analyses to help determine compliance with § 25.1309(b), the following descriptions of the probability terms used in this regulation and this AC have become commonly-accepted as aids to engineering judgment. They are usually expressed in terms of acceptable numerical probability ranges for each flight-hour, based on a flight of mean duration for the airplane type. However, for a function which is used only during a specific flight operation; e.g., takeoff, landing, etc., the acceptable probability should be based on, and expressed in terms of, the flight operation's actual duration.

(1) Probable failure conditions are those having a probability greater than on the order of 1 X 10^{-5}.

(2) Improbable failure conditions are those having a probability on the order of 1 X 10^{-5} or less, but greater than on the order of 1 X 10^{-9}.

(3) Extremely Improbable failure conditions are those having a probability on the order of 1 X 10^{-9} or less.

This is the FAR itself ("law", not advice), but it uses non-quantitative terms

This is the guidance (not "law", but used to interpret law's intent) which **is** in quantitative terms.

Figure 8: Excerpt from AC 25.1309-1A

4.4 Does a Reliability Requirement Inform a Quality Parameter Choice?

ARP9013 requires that the Quality Parameter is selected with knowledge about the consequences of nonconforming product (section 3.3.3). Example: An OEM who is responsible for their product's reliability may say that this "knowledgeable individual" should be one of their design staff, perhaps someone like a liaison engineer. If these "consequences" include product reliability, then this equation is offered:

Prob (it works) = [Prob (it works given it conforms) × Prob (it conforms)] +
[Prob (it works given it does not conform) × Prob (it does not conform)]

Figure 9: The "Law of Total Probability" in Reliability Engineering

- Operational reliability requirements go on the left, in Prob (it works).
- Design reliability values (from accelerated life testing, or models built from

test results) go on the right, in Prob (it works given it conforms).
- The value for Prob (it works given it does not conform) may be estimated as a worst-case from field experience data, or may be conservatively set at zero.
- Prob (it does not conform) = 1 – Prob (it conforms).
- With those values, the equation can be solved for the worst-case value for Prob (it conforms), which is the Initial Reliability Requirement.

5. Silent Specs, Costs, and Reliability

For new processes, users might like a table of default quality parameter values:

- Many drawings and specs are silent about the quality parameter(s) to use; some control technologies that have only been developed since the mil-specs were cancelled.
- Reliability is often calculated on the assumption of quality levels similar to past production.
- Costs are difficult to estimate more precisely than the order of magnitude.

With at least a default quality parameter for everything, the aerospace industry could comply with the requirement to "ensure" and to "determine" conformance to every requirement. ARP9013 Figure B1 provides a table that could be used, or serve as a starting point for, this default.

6. Summary

- It is of interest to a commercial airplane manufacturer to be able to PROVE that they have "ensured" and "determined" the conformance of every part of every airplane.
- By satisfying the quality parameter, that manufacturer can "ensure" and "determine" that conformance – IF the quality parameter has been "selected or approved by individuals who understand the consequences of nonconformance".
- If the Type Design has no quality parameter in it, the various Operations customers have three courses of action to choose among:
    - Operations may act on the basis that the legacy values are good enough, and work to standards similar to what were in the military specifications 1942-1996. See ARP9013 Figure B1 for illustration.
    - Operations may act solely on the basis of cost-optimization. The rationale here could be that if there was any safety concern, the Type Design would have either required 100% inspection or given a considered quality parameter.
    - Operations may contact the Type Design owners to calculate a reliability-based quality parameter, or else to validate one proposed from Operations. FAA-recommended reliability values such as in AC 25.1309-1A are a demanding but credible basis for quality protection level selection.

This gives the Statistical Engineer the starting point for probabilistic design of the entire production system.

Acknowledgements

Credit goes to my Boeing colleagues and to the SAE authors of ARP9013 for providing the opportunity to connect these important concepts in Statistical Engineering.

References

“CRE Primer”, Quality Council of Indiana, 2009
FAA Advisory Circular 21-43
FAA Advisory Circular 25.1309-1A
Federal Aviation Regulations 21.143 and 21.165
FAA “Best Practice / Statistical Quality Control”, 2001 and 2006 editions
MIL-HDBK-217, “Reliability Prediction of Electronic Equipment”
MIL-HDBK-1916, “Companion Document to MIL-STD-1916”
MIL-STD-756, “Reliability Modeling and Prediction”
“Out of the Crisis”, W. Edwards Deming, 1987
SAE ARP9013, “Statistical Methods of Product Acceptance”, 2005