

## 講義の試験と製品検査

メタデータ	言語: jpn 出版者: 公開日: 2014-09-02 キーワード (Ja): キーワード (En): 作成者: 熊谷, 正朗 メールアドレス: 所属:
URL	<a href="https://tohoku-gakuin.repo.nii.ac.jp/records/204">https://tohoku-gakuin.repo.nii.ac.jp/records/204</a>

## 講義の試験と製品検査

学期末が近づくと、学生さんの成績評価で頭が痛くなってきます。おそらく、多くの大学教員の大きな悩み事の一つであろうかと思いません。

大学における試験は、入学試験と、在学生の成績評価のための試験に大別され、その評価方法は大きく異なります。一般的には、入学試験は合格者数に達するまでの相対的な順位付けを行うことが主目的で、知識や判断力の絶対的なレベルの判定ではありません。そのため、試験問題はいかに受験生の能力を反映し、力の差が点数の差として現れるようにするかが重要になっていると考えられます。

これに対して、在学生の成績は、所定のレベルに達したかを判断します。そのレベルは、厳密には、事前にシラバス(授業概要を提示したもの)などで提示されることになっています。また、大学においては60点を合格点とし、60点以上なら合格して「単位」と呼ばれる数値を得られ、59点以下は不合格として単位にはなりません。大学の卒業や進級がこの単位によって決まるため、「レベルに達したらちょうど60点」となるような評価方法をつくらなければならない、この点については入試よりも面倒です。

世の中のさまざまな分野での試験や評価にも、この2種類があります。なにかを導入しようというときには、性能やコストの面で複数

を比較し、優れるものを採用しますし、製品の出荷検査などでは合格と認められる性能を確保できているかを確認します。

選定のための評価は多面的な評価をいかに組み合わせるか、という点が重要になります。たとえば、導入コスト、処理速度のような性能、将来の拡張性、技術的なリスクの点で評価しようとしてみます。まずは、各々共通の数値化(点数化)を行い、単位をそろえる必要があります。しかし、技術として枯れているからリスクが低そう、新しいから未知のトラブルがありそう、といったところは点数化がむずかしいところです。

点数化できたとしても、候補Aは導入コストは安いけど(現時点の要求は満たしていても)将来的な拡張の余地は少ない、候補Bはコストは高いけど発展の余地あり、という場合には、今の予算を考えるか、将来性を考えるかという判断の分かれ目があります。カンで判断することも一つですが、その場合は客観性に疑問符がつき、後々説明がむずかしくなります。そのため、たとえば、

$$\text{導入コスト点数} \times 0.4 + \text{性能} \times 0.2 \\ + \text{拡張性} \times 0.2 + \text{リスク} \times 0.2$$

というように、点数の重み付け和を求めます。この $\times 0.?$ の部分がどれを重視するかにあたります。入試では、科目ごとの配点という形で

**熊谷正朗**—KUMAGAI MASAOKI—

東北学院大学 工学部 機械知能工学科 教授

東北学院大学工学部 教授／仙台市地域連携フェロー(ロボットメカトロ系担当)。2000年東北大学大学院工学研究科修了、博士(工学)。同大助手を経て、03年東北学院大学講師、助教授、准教授を経て、13年より教授。ロボメカ系開発を専門とし、メカの設計からマイコンやサーバのソフト開発までを行う。「基礎からのメカトロニクス講座」や地域企業訪問も実施中。



これが現れます(工学系は数学と理科の点数が大きいなど)。ただし、この係数は評価の前に決めておくべきです。個別の評価が出てから係数を決めるのでは、恣意的な調整ができてしまうので、あえてこの方法を採用の意味が減ってしまいます。

一方、製品の良否判定は、工場などでは重要な要素といえ、昨今ではメカトロ技術の進展とともに検査の自動化も進み、製造ライン上で一括して行う検査も増えています。これらには機能的な試験の他、カメラなどを用いた外観検査などがあります。ここで問題になるのは、評価の誤差です。評価が正確なら、合格点を明確に定めることが容易ですが、評価手法そのものにばらつきがあると、そうはいきません。評価の正誤には以下の4パターンがあります。

- (A) 製品は良 判定も良
- (B) 製品は良 判定は不良
- (C) 製品は不良 判定は良
- (D) 製品は不良 判定も不良

いうまでもなく、(A) (D)が望む結果です(生産の観点ではDは好ましくありませんが)。問題は(B)と(C)です。一見するとどちらも誤判定ですが、両者は大きく意味が違います。製品という観点からは、(C)では不良品が出荷されてしまうという意味で大問題です。一方、(B)が多いと本来は出荷できるものを捨ててしまう

ことになります。評価のばらつきが大きいと(B) (C)が増えますが、(C)を避けるには判断のレベルを十分に上げ、ちょっとでも疑わしかったら不良扱いとする(B)にしなければなりません。その結果、歩留まりの低下＝生産コストの上昇になります。この観点から、不良品を出さない製造が重要であると同時に、評価方法が非常に重要であることがわかります。最近では画像処理技術の進歩にともない、カメラを用いた検査などが多くなっていますが、いまだ人間並みの判断力はありません。光の当て方などの撮影条件の最適化が、判定をより正確にする鍵の一つです。

さて、大学の試験は、というと、先ほど述べたように卒業や進級に関わるため、評価のばらつきによって本来合格である学生さんに不合格をつけるわけにはいきません。つまり、上の分類でいえば(B)を避ける必要があって(C)は許容、という工場の出荷検査と逆のパターンになります。かといって必要以上に緩くするわけにもいかないので、どう判定をつけるか、頭が痛くなるわけです。そうですね、講義一つやる上での苦勞の半分以上は採点作業といっても過言ではないかもしれません。

■参考キーワード：偽陽性(False positive)、偽陰性(False negative)