

読み取り誤差の結果について

2020.9.23 吉村洋介

今年度は読み取り誤差の実験を Web 上のクイズ

http://kuchem.kyoto-u.ac.jp/ubung/yyosuke/uebung3/basics/readingscale_2020.htm

の形で行いました。皆さんに提出いただいた結果を取りまとめた結果を紹介します。

1. 基礎データ

提出いただいた excel シートは全部で 47。そこに記載された 1904 個のデータから、あまりに外れたもの (± 0.1 以上) を除く 1839 個のデータについて統計を取っています。

出題の設定値の小数部の出現頻度は図 1 のとおり。最初に確認の意味もあって、0.0 から始めているので、0.0 となる設定が多くなっていますが、それ以外はおおむね 180 回程度ずつばらつきなく発生されています。

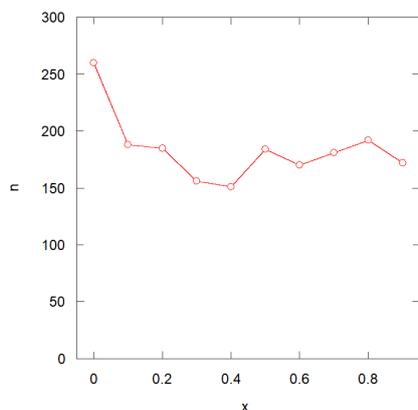


図 1 出題される設定値小数部 x の出現頻度。必ず最初に 0.0 から始めるので、0 が多くなっているが、ほぼ均等に発生されている。

1. 偏倚の出現頻度

図 2 に出題の設定値の小数部の値 x ごとに、マウスでクリックした位置の小数部の値 x_p の偏倚 $\Delta x = x_p - x$ の出現頻度をプロットした結果を示します。

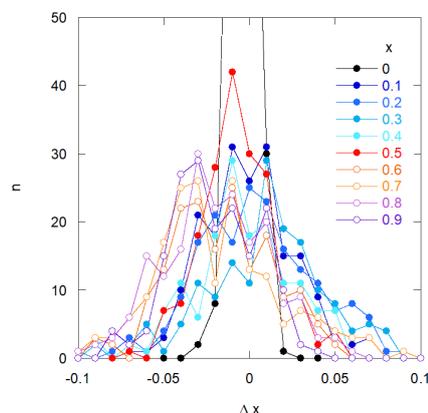
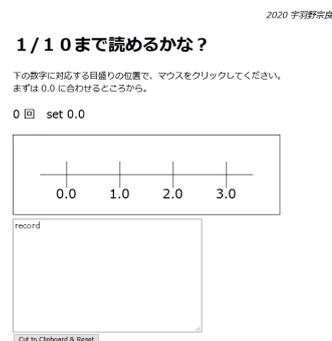


図 2 設定値小数部 x ごとのマウスでクリックした位置 x_p の偏倚 $\Delta x = x_p - x$ の出現頻度 $n(\Delta x)$ 。 $x = 0$ の $\Delta x(-0.01) = 109$, $\Delta x(0.00) = 110$, $\Delta x(0.01) = 30$ 。

さすがに $x = 0.0$ はよく一致していますが、若干低い側に合わせる傾向があるようです。 $x < 0.5$ については $\Delta x > 0$ の方に、 $x > 0.5$ については $\Delta x < 0$ の方に少し分布が偏ります。これらを総括して、設定値の小数部の値によらず偏倚の出現頻度をまとめたのが図 3 です。



± 0.05 の幅に入ってくるのは 1839 件中 1642 件。目視で 1/10 まで 89 % は正しく設定されている、あるいは 11 % は正しく設定できていないことになります。またピークの位置が $\Delta x = 0$ より少し負の側に来るようです。

これを各設定値の小数部の値についてみたものが図 4 になります。 $x = 0.0$ と 0.5 で設定精度が高くなるのが分かります。ただし $x = 0.5$ について対称的な挙動ではなく、 x が大きいほど設定値が外れる傾向にあります。

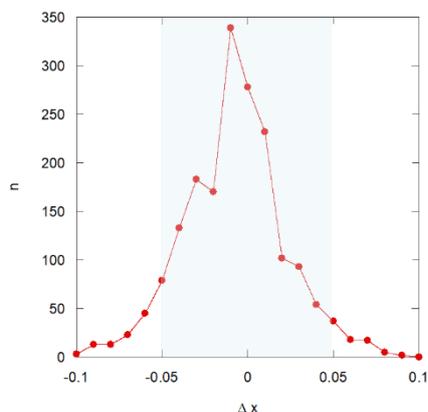


図 3 設定偏倚 Δx 出現頻度。
 ± 0.05 (図中薄青の部分) に
89 % が集まる。

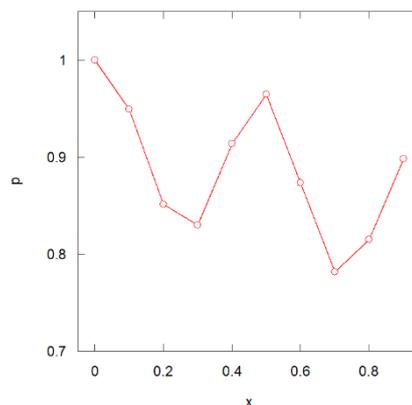


図 4 設定値小数部 x ごとの、
設定偏倚 Δx が ± 0.05 に収まる
確率。

これをそれぞれの x での偏倚の平均で表現したものが図 5 です。図中のバーの長さは各 x における標準偏差に対応しています。全偏倚の平均 $\langle \Delta x \rangle$ は -0.008 で、ほぼ $x = 0.5$ における偏倚値の平均に等しくなっています。

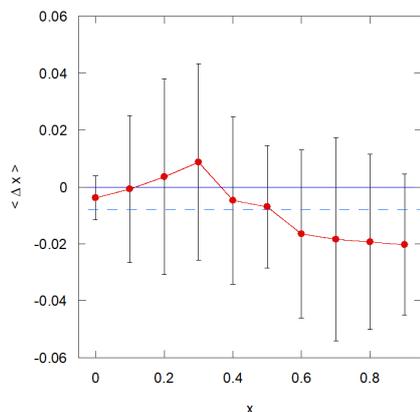


図 5 設定値小数部 x ごとの、
設定偏倚の平均値 $\langle \Delta x \rangle$ 。バーの
長さは標準偏差に対応。薄青の
破線は全体の平均値 -0.0078 。

従来の物差しにカーソルを合わせるタイプの実験では、ほぼ 0.5 を中心に対称な結果が得られていたのですが、今回の web を利用した読み取り誤差の実験では、設定値が全体として小さい方にずれている傾向が認められます。これはディスプレイ上のマウスポインターの位置が、若干右方に認識されている (ほぼ 1 ピクセル) ためと思われます。現在採用している標準のマウスポインターの形状(⇨)など、含めてよく検討する必要があるようです。