Matematicko-statistické zpracování výsledků smyslového hodnocení jakosti piva

Jiří Cuřín, Pokusné a vývojové středisko, oborové ředitelství Pivovary a sladovny, Praha

663.439.1
543.92
519.2

Vzhledem k tomu, že jde o nový způsob vyhodnocení výsledků degustace, uveřejňuje redakce uvedený článek. Je zapotřebí však zdůraznit, že takovéto matematické zkoumání lze dělat jen v těch případech, když degustaci provádějí školení degustátoři.

Hodnocení jakosti všech potravinářských výrobků, tedy i piva, se neobejde bez organoleptického posuzování. Vztah mezi chemickými a fyzikálními objektivně měřitelnými vlastnostmi a smyslově vnímanými podněty není ještě natolik znám, aby bylo možno jednoznačně posoudit tyto podněty pouze z chemických či fyzikálních stanovení. Největší význam mají podněty vnímané chutí a čichem, úhrnně nazývané chutností, které jsou jedním z rozhodujících činitelů ovlivňujících přístup spotřebitele k výrobku. Smyslové hodnocení má i výhody v rychlosti, menší pracnosti i náročnosti na laboratorní vybavení. Z těchto důvodů je smyslové hodnocení již odedávna běžně používanou potravinářskou zkušební metodou.

Jako u každé analytické metody je zapotřebí i při smyslovém posuzování dodržovat jistá pravidla a postupy, aby bylo možno ze získaných výsledků vyvodit správné závěry. Prvním předpokladem je především podržet základní podmínky při vlastním hodnocení. Těmito problémy se zabývalo několik autorů [1, 2, 3, 4].

Vlastní hodnocení v pivovarském oboru se podstatně zlepšuje, i když lze provádění zkoušek ještě vytknout některé chyby. Nesporně velkým úspěchem je již hodnotitelská komise v každém závodě, jejíž členové prošli organoleptickými testy. Současné hodnocení je zaměřeno v souhlase s požadavky nového způsobu řízení především k promítnutí úrovně jakosti do odměňování pracovníků.

S rozvíjením nového způsobu řízení se úroveň jakosti bude nutně projevovat i druhou ekonomickou formou, tj. přímo ve velikosti odbytu. Vzhledem k tomu, že pivo je u nás již po řadu let nedostatkovým výrobkem a odbyt je skoro předem zajištěn, není zatím všem detailům chuťových vlastností věnována odpovídající péče. Odbyt piva je však již tak vysoký, že postupně se musí trh nasytit. A v této chvíli otázka chuti, resp. přesněji chutnosti, opět vyvstane v celé své závažnosti.

Výsledky smyslového hodnocení nejsou samy o sobě dostatečným podkladem pro zodpovědné provedení zásahů ve výrobě. Je nutno ještě znát věrohodnost výsledků, abychom své názory měli dostatečně podloženy. Tuto informaci mohou poskytnout pouze údaje zpracované matematicko-statistickými metodami. O nutnosti použití těchto metod hovoří skoro všichni autoři, zabývající se smyslovým hodnocením potravin. V současné době se u nás zpracování výsledků většinou omezuje na vypočtení aritmetického průměru. Posouzení pouze podle průměru však často vede k špatným závěrům. Zvláštní význam má tato okolnost ve výzkumu jakéhokoli charakteru. Zde se velmi často dostáváme do paradoxní situace. Zatímco velmi přesně analyzujeme různé dílčí fyzikální a chemické vlastnosti, při posouzení chutnosti se spokojíme stručným konstatováním dosaženého průměru hodnocení. Z těchto důvodů je zapotřebí pracovníky oboru upozornit na možnosti, které v organoleptice a samozřejmě i v celé laboratorní práci skýtá aplikace matematicko-statistických metod.

Běžně u nás používaným způsobem hodnocení jakosti je posuzování podle bodovacího systému. Výhodou tohoto postupu je kromě jiného možnost kvalitativního vyjádření jednotlivých složek chutnosti, nevýhodou je vznik chyb rozdílným výkladem kritérií jednotlivými hodnotiteli. V podstatě jde o porovnávací hodnocení buď se srovnávacím vzorkem, nebo na základě vjemů uchovaných v paměti. Možnosti použít srovnávacího vzorku se zatím u nás nevyužívá. Kromě standardizace celého hodnocení je podání srovnávacího vzorku před hodnocením žádoucí i z fyziologického hlediska, poněvadž zajišťuje přípravu organismu k posuzování.

Bodovací systém může mít buď rovnoměrné rozdělení bodů podle jednotlivých složek, nebo rozdělení může vyjadřovat jejich důležitost. V prvním případě je nutno použít pro konečné vyhodnocení tzv. koeficientů důležitosti. Vyjadřují podíl významu složek na úrovni celkové jakosti. Tento způsob se přes jistou větší složitost zdá pro široké upotřebení výhodnější, poněvadž vysoké bodové rozdíly mohou vést u některých hodnotitelů k jisté, mimovolné snaze po nivelizaci výsledků.

V praxi jsou u nás pro hodnocení piva uplatňovány dva systémy, a to stobodový a sedmdesátipětibodový. Stobodový systém je určen především pro spojení úrovně jakosti s odměňováním pracovníků a zahrnuje jak smyslové, tak i laboratorní posouzení výrobku. Tomuto účelu vcelku dobře vyhovuje. Pro detailnější posouzení chuti, nutné pro nejrůznější technologické zkoušky, se však nehodí, poněvadž veškerý chuťový i čichový vjem soustřeďuje pouze do jediného kritéria. Pro přesnější zhodnocení je výhodnější sedmdesátipětibodový systém VÚPS, hodnotící odděleně chuť i vůni, hořkost a dojem po napití. Avšak ani tento systém nedosahuje poměrně hlubokého členění vjemů užívaného v zahraničí. Smyslem všech těchto systémů je ponechat hodnotiteli co nejméně libovůle a zajistit tak přesnost hodnocení. Tak např. *Gray* [4] rozlišuje v chuti piva celkem šest složek. Je to všeobecný dojem, hořkost, trvání hořké chuti v ústech, chmelové aróma, plnost a cizí příchuti. Každou z těchto složek ještě dále dělí a přesně určuje.

Základní podmínkou pro matematicko-statistické zpracování výsledků je individuální hodnocení všemi posuzovateli [5]. Kolektivní posouzení nejen že nikdy neposkytne správné výsledky, ale neumožňuje ani jakékoli další zpracování údajů. Nemůžeme totiž na něj aplikovat matematické modely rozdělení tzv. náhodné veličiny, které jsou základem všech matematicko-statistických postupů. Kolektivní hodnocení komise bývá vždy ovlivněno názorem jednotlivce, který může být dobrý, avšak i špatný. Z tohoto hlediska se jeví ustanovení platného stobodového systému o kolektivním hodnocení pěnivosti i průzračnosti jako nesprávné. Potvrzují to konečné výsledky získané v oborových přehlídkách jakosti lahvového piva pořádaných OŘ. Pro ověření objektivity hodnocení průzračnosti bylo provedeno vedle smyslového hodnocení i objektivní změření Sigristovým fotometrem. Například z počtu 72 hodnocených vzorků nevyhovělo podle měření fotometrem 11 vzorků. Přitom však hodnotitelé neoznačili žádný vzorek jako nevyhovující a pouze 5 z těchto vzorků ohodnotili sníženým počtem bodů. Stejná situace je i v pěnivosti, která je posuzována vůbec jenom jednou. Když byl hodnotitelům v sérii 5 vzorků předložen tentýž vzorek dvakrát, pak v 7 ze 13 případů hodnotili pěnivost rozdílným stupněm. Rozdíly v průměrech chuti a vůně, hodnocené individuálně, byly podstatně nižší. Kromě vlivu komise je zde ovšem nutno vzít v úvahu i variabilitu pěnivosti v různých lahvích téhož vzorku. V každém případě však výsledky, v souhlase s teoretickými předpoklady, mluví pro vícenásobné individuální hod[illegible]ení.

Výsledky z[illegible]ané bodovacím systémem se z hlediska matema[illegible]ko-statistického zpracovávají poměrně nejobtíž[illegible]i. Podstatou každého takového zpracování je z[illegible]ožnění daného hodnocení (pro matematicko-statistické účely je to soubor údajů) s vhodným matematickým modelem a využití jeho vlastností pro získání žádaných informací. Jde o modely rozdělení náhodné veličiny a v praxi se jich několik používá. Pro účely zpracování výsledků organoleptických hodnocení je nejdůležitější vedle normálního rozdělení tzv. rozdělení t nebo Studentovo a rozdělení F nebo Snedecorovo. Volba modelu závisí jednak na vlastnostech hodnoceného souboru, jednak na jeho rozsahu. Je proto třeba odlišně zpracovávat soubory velkého a malého rozsahu. Potřebné údaje jsou pro všechny modely tabelovány. Níže jsou uvedeny pouze základní důležité skutečnosti a pro podrobnější seznámení lze doporučit některou z učebnic a příruček [7, 8, 9, 10, 11], kde jsou uvedeny i potřebné tabulky hodnot.

Základní veličinou, s níž budeme operovat, je aritmetický průměr definován jako

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

kde

n je počet hodnot; $\sum_{i=1}^{n} x_i$ — součet hodnot;

Další veličinou je rozptyl čili průměr čtverců odchylek od aritmetického průměru

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

a konečně směrodatná odchylka, která je druhou odmocninou z rozptylu

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Již samotný rozptyl a zvláště směrodatná odchylka ve srovnání s průměrem dávají představu o vyrovnanosti výsledků. Můžeme stanovit i interval, ve kterém se se zvolenou pravděpodobností či jistotou bude pohybovat skutečná hodnota jakosti. Pravděpodobnost se uvádí buď jako desetinné číslo, nebo v procentech. Tedy pravděpodobnost 0,95 je totéž, jako pravděpodobnost 95%. Nejčastěji se používá pravděpodobnosti 0,95 nebo 0,99.

V dalším budou údaje o jakosti, získané od jednotlivých hodnotitelů při hodnocení téhož vzorku, nazývané údaji hodnocení. Pro posouzení výsledků malého počtu údajů (tj. pod 30) se používá jako matematického modelu t rozdělení náhodné veličiny. Platí zde, že skutečná hodnota jakosti bude ležet se zvolenou pravděpodobností $(1-p)$ v intervalu

$$\left[\bar{x} - t_p \frac{s}{\sqrt{n-1}} \; ; \; \bar{x} + t_p \frac{s}{\sqrt{n-1}}\right],$$

kde

$\bar{x}$ je aritmetický průměr hodnocení

n — počet údajů hodnocení;

t_p — tabelovaná hodnota rozdělení t pro zvolenou pravděpodobnost p a počet stupňů volnosti $\nu = n-1$;

s — směrodatná odchylka.

V tomto případě totiž uvažujeme v podstatě o tom, s jakou pravděpodobností p veličina přesáhne dané meze, čili výraz $(1-p)$ udává pravděpodobnost či jistotu, s jakou v daném intervalu bude ležet. Počtem stupňů volnosti se rozumí počet veličin, které je možno v daném matematickém modelu volit libovolně. Má-li např. 10 veličin mít aritmetický průměr 5, můžeme 9 z nich zvolit libovolně a desátá je již určena volbou průměru, čili počet stupňů volnosti je v tom případě 9. Pro hodnocení nad 30 údajů se používá normálního rozdělení a skutečná hodnota jakosti leží s pravděpodobností p v intervalu

$$\left[\bar{x} - u_p \frac{s}{\sqrt{n-1}} \; ; \; \bar{x} + u_p \frac{s}{\sqrt{n-1}}\right],$$

kde

$\bar{x}$ je aritmetický průměr hodnocení;

n — počet údajů hodnocení;

s — směrodatná odchylka;

u_p — tabelovaná hodnota u normálního rozdělení pro zvolenou pravděpodobnost p.

Hodnota u je pro pravděpodobnost 0,95 $u_{0,95} = 1{,}96$, pro pravděpodobnost 0,99, $u_{0,99} = 2{,}58$.

Vedle tohoto rozboru pro utvoření správného závěru je zapotřebí ještě velmi často posoudit, zda rozdíl průměrného hodnocení dvou vzorků je, či není statisticky významný. Pro tento účel považujeme získané údaje obou hodnocení za výběrové soubory a zkoumáme, zda můžeme předpokládat, že pocházejí ze stejného rozsáhlejšího souboru základního, čili že průměry obou původních souborů jsou si rovny ($\mu_1=\mu_2$). Vytváříme tzv. nulovou hypotézu, kterou buď příjmeme nebo zamítneme. Musíme rozeznávat případ, kdy obě hodnocení jsou na sobě nezávislá (hodnoceno různými komisemi) nebo kdy mezi nimi existuje jistá spojitost (hodnoceno současně stejnou komisí).

V obou případech nejprve musíme zvolit hladinu významnosti, tj. pravděpodobnost ($1-p$), se kterou chceme pracovat. Nejčastěji se pracuje s pravděpodobností 0,95 ($p=0{,}05$) nebo 0,99 ($p=0{,}01$).

I. V případě nezávislých hodnocení musíme nejprve zjistit, zda mezi rozptyly obou hodnocení je významný rozdíl. K tomu použijeme testování pomocí rozdělení F. Vypočteme odhad směrodatné odchylky původních souborů $\hat{\sigma}_1$ a $\hat{\sigma}_2$, tj. $\hat{\sigma}_1$ a $\hat{\sigma}_2$ přitom

$$\hat{\sigma} = s\sqrt{\frac{n}{n-1}},$$

kde

s je směrodatná odchylka hodnocení,
n — počet údajů hodnocení

a pak zjistíme, zda platí $\sigma_1^2=\sigma_2^2$ (nulová hypotéza).

Vypočteme

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

$\hat{\sigma}_1^2$ má počet stupňů volnosti $\nu_1=n_1-1$ a $\hat{\sigma}_2^2$ má počet stupňů volnosti $\nu_2=n_2-1$; $\hat{\sigma}_1^2$ a $\hat{\sigma}_2^2$ volíme tak, aby $F\geqq 1$. Z tabulky zjistíme hodnotu $F_{\frac{p}{2}}$ pro počet stupňů volnosti ν_1 a ν_2 a zvolené p.

V případě že vypočtené $F>F_{\frac{p}{2}}$, pak rozdíl rozptylů obou hodnocení je nutno považovat za významný, čili že $\sigma_1^2\neq\sigma_2^2$ a nulovou hypotézu zamítáme. V opačném případě $\sigma_1^2=\sigma$.

Nyní přistoupíme k posouzení významnosti rozdílu obou průměrů. Testováním podle t rozdělení zjistíme, zda průměry obou původních souborů jsou stejné ($\mu_1=\mu_2$). Postup volíme podle zjištěného vztahu mezi σ_1^2 a σ_2^2.

1. $\sigma_1^2=\sigma_2^2$

Vypočteme

$$t = \frac{|\bar{x}_1-\bar{x}_2|}{\sqrt{n_1s_1^2+n_2s_2^2}} \cdot \sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}$$

kde

$\bar{x}$ je aritmetický průměr hodnocení;
n — počet údajů hodnocení;
s — směrodatná odchylka hodnocení.

Vypočtenou hodnotu t porovnáme s tabelovanou hodnotou t_p pro zvolené p a počet stupňů volnosti $\nu=n_1+n_2-2$. V případě, že $t>t_p$ pak $\mu_1\neq\mu_2$, nulovou hypotézu zamítáme a rozdíl mezi průměry hodnocení musíme považovat za statisticky významný.

2. $\sigma_1^2\neq\sigma_2^2$

Vypočteme

$$t = \frac{|\bar{x}_1-\bar{x}_2|}{\sqrt{\dfrac{s_1^2}{n_1-1}+\dfrac{s_2^2}{n_2-1}}}$$

a srovnáme ho s

$$t_p^* = \frac{t'_p\dfrac{s_1^2}{n_1-1}+t''_p\dfrac{s_2^2}{n_2-1}}{\dfrac{s_1^2}{n_1-1}+\dfrac{s_2^2}{n_2-1}},$$

kde

t'_p je tabelovaná hodnota t pro p a počet stupňů volnosti $\nu_1=n_1-1$;
t''_p — tabelovaná hodnota pro p a počet stupňů volnosti $\nu_2=n_2-1$;
ostatní symboly jako v případě 1.

Jestliže $t > t_p^*$ pak $\mu_1\neq\mu_2$, nulovou hypotézu opět zamítáme a rozdíl mezi průměry musíme považovat za statisticky významný.

II. V případě závislých hodnocení, který je ostatně nejčastějším, je posouzení jednodušší. Pro testování opět použijeme t rozdělení. U všech párových hodnot vypočteme rozdíl

$$d_i = x_{1i} - x_{2i},$$

kde x_{1i} a x_{2i} jsou párové údaje o jakosti obou vzorků od jednoho hodnotitele.

Dále vypočteme

$$t = \frac{|\bar{d}|\sqrt{n-1}}{s_d},$$

kde

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

$$s_d = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i-\bar{d})^2}$$

n je počet párů.

Vypočtené t srovnáme s t_p pro zvolené p a $\nu = n-1$ stupňů volnosti.

Příklad:

Máme k dispozici organoleptické hodnocení 4 vzorků A, B, C a D. Vzorky A, B, C byly hodnoceny současně stejnou komisí, vzorek D hodnotila jiná komise. Chceme vědět, zda můžeme vzorky B, C a D považovat za totožné se srovnávacím vzorkem A. Výsledky degustace jsou uvedeny v *tabulce 1.* U všech čtyř hodnocení je vypočten aritmetický průměr $\bar{x}$, rozptyl s^2, směrodatná odchylka s a interval spolehlivosti, v němž se jakost vzorků bude pohybovat s pravděpodobností 95 % (0,95). Pro výpočet významnosti rozdílů průměru hodnocení nejprve zvolíme hladinu významnosti $p = 0{,}05$. Posuzujeme tedy, zda se průměry liší s pravděpodobností ($1-p$), tj. 0,95. Poněvadž vzorky A, B, C byly hodnoceny stejnou komisí a je tedy mezi výsledky určitý vztah, užijeme postupu podle odstavce II. Hodnoty d_i, $d_i-\bar{d}$, a $(d_i-\bar{d})^2$ jsou uvedeny rovněž v *tabulce 1.*

a) *Vztah A — B*

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{+9}{9} = +1$$

$$s_d = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i-\bar{d})^2} = \sqrt{\frac{164}{9}} = 4{,}27$$

$$t = \frac{|\bar{d}|\sqrt{n-1}}{s_d} = \frac{|+1|\sqrt{8}}{4{,}27} = 0{,}66$$

Počet stupňů volnosti $\nu = n-1 = 9-1 = 8$, takže $t_{0{,}05} = 2{,}31$ čili $t_{0{,}05} > t$ z čehož plyne, že rozdíl

Tabulka 1

Hodnocení A			Hodnocení B			Hodnocení C			Hodnocení D		
Body	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	Body	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	Body	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	Body	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
71	+ 0,89	0,7921	68	− 1,11	1,2321	72	+ 0,44	0,1936	72	+ 2,87	8,2369
71	+ 0,89	0,7921	69	− 0,11	1,0121	72	+ 0,44	0,1936	71	+ 1,87	3,4969
65	− 5,11	26,1121	72	+ 2,89	8,3521	67	− 4,56	20,7936	73	+ 3,87	14,9769
74	+ 3,89	15,1321	71	+ 1,89	3,5721	75	+ 3,44	11,8336	69	+ 0,13	0,0169
64	− 6,11	37,3321	70	+ 0,89	0,7921	66	− 5,56	30,9136	71	+ 1,87	3,4969
71	+ 0,89	0,7921	70	+ 0,89	0,7921	72	+ 0,44	0,1936	69	+ 0,13	0,0169
72	+ 1,89	3,5721	67	− 2,11	4,4521	73	+ 1,44	2,0736	65	− 4,13	17,0569
72	+ 1,89	3,5721	70	+ 0,89	0,7921	74	+ 2,44	5,9536	63	− 6,13	37,5769
71	+ 0,89	0,7921	65	− 4,11	16,8921	73	+ 1,44	2,0736	—	—	—
Σ 631		88,8889	622		36,8889	644		74,2224	553		84,8752
$\bar{x}$ = 70,11 s^2 = 9,8765 s = 3,14			$\bar{x}$ = 69,11 s^2 = 4,0988 s = 2,02			$\bar{x}$ = 71,56 s^2 = 8,2469 s = 2,87			$\bar{x}$ = 69,13 s^2 = 10,6094 s = 3,26		
Interval spolehlivosti při $1 - p = 0,95$ 70,11 ± 2,56			69,11 ± 1,65			71,56 ± 2,34			69,13 ± 2,91		
Rozpětí	10		7			9			—		

Součet rozpětí Λ = 26

Tabulka 1 (pokračování)

Vztah A—B			Vztah A—C		
d_i	$d_i - \bar{d}$	$(d_i - \bar{d})^2$	d_i	$d_i - \bar{d}$	$(d - \bar{d})^2$
+ 3	+ 2	4	− 1	+ 0,44	0,1936
+ 2	+ 1	1	− 1	+ 0,44	0,1936
− 7	− 8	64	− 2	− 0,56	0,3136
+ 3	+ 2	4	− 1	+ 0,44	0,1936
− 6	− 7	49	− 2	− 0,56	0,3136
+ 1	0	0	− 1	+ 0,44	0,1936
+ 5	+ 4	16	− 1	+ 0,44	0,1936
+ 2	+ 1	1	− 2	− 0,56	0,3136
+ 6	+ 5	25	− 2	− 0,56	0,3136
Σ + 9		164	− 13		2,2224
$\bar{d}$ = + 1,00			$\bar{d}$ = − 1,44		

mezi průměry hodnocení nemůžeme považovat za statisticky významný.

b) *Vztah A — C*

$$\bar{d} = \frac{-13}{9} = -1{,}44$$

$$sd = \sqrt{\frac{2{,}2224}{9}} = 0{,}50$$

$$t = \frac{|-1{,}44|\ \sqrt{8}}{0{,}50} = 8{,}15$$

Počet stupňů volnosti $\nu = 9 - 1 = 8$, $t_{0,05} = 2{,}31$, tudíž $t_{0,05} < t$ z čehož plyne, že rozdíl mezi průměry hodnocení je statisticky významný. Srovnání obou případů (a, b) ukazuje že skoro stejný rozdíl průměrů hodnocení neznamená ještě stejný vztah a konečné závěry ovlivňuje i rozptyl hodnocení.

c) *Vztah A — D*

Při hodnocení souvislosti mezi A a D musíme použít způsobu podle odst. I. Jde o vzorky hodnocené dvěma různými komisemi. Nejprve musíme proto posoudit rozptyly obou hodnocení.

$$\hat{\sigma}_1^2 = \frac{s_1^2\, n_1}{n_1 - 1} = \frac{10{,}6094 \,.\, 8}{7} = 12{,}13$$

$$\hat{\sigma}_2^2 = \frac{s_2^2\, n_2}{n_2 - 1} = \frac{9{,}8765 \,.\, 9}{8} = 11{,}11$$

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{12{,}13}{11{,}11} = 1{,}09$$

Počet stupňů volnosti $\nu_1 = n_1 - 1 = 7$, $\nu_2 = n_2 - 1 = 8$, takže $F_{0,025} = 4{,}53$. Protože platí, že $F_{0,025} > F$, můžeme považovat $\sigma_1^2 = \sigma_2^2$ a užijeme postupu I/1.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} =$$

$$= \frac{|70{,}11 - 69{,}13|}{\sqrt{9 \,.\, 9{,}8765 + 8 \,.\, 10{,}6094}} \cdot \sqrt{\frac{(9 \,.\, 8\,(9 + 8 - 2)}{9 + 8}} = 0{,}59$$

Počet stupňů volnosti $\nu = n_1 + n_2 - 2 = 15$ takže $t_{0,05} = 2{,}13$ $t_{0,05} > t$, čili rozdíl průměrů hodnocení nemůžeme považovat za statisticky významný.

Kromě uvedených postupů lze pro méně přesné posouzení rozdílů v hodnocení jakosti použít i jednoduché metody mnohonásobného porovnávání vypracované *Tukeyem* [12]. Pro její použití je nutné, aby u všech hodnocení byl stejný počet údajů. Metoda vychází z porovnání rozpětí, tj. rozdílu mezi nejvyšším a nejnižším údajem hodnocení. V *tabulce 2* jsou uvedeny koeficienty, jimiž pro 95% nebo 99% pravděpodobnosti násobíme součet rozpětí všech srovnávaných hodnocení. Rozdíly mezi hodnoceními považujeme za významné, jestliže rozdíly součtu všech údajů hodnocení jsou větší než tato hodnota.

Příklad:

U hodnocení A, B, C v *tabulce 1* je rozpětí 10, 7 a 9. Součet rozpětí je 10 + 7 + 9 = 26. Pro tři skupiny, devět opakování a 95% pravděpodobnost

Tabulka 2
Koeficienty pro výpočet spolehlivosti rozdílů v jakosti Tukeyovou metodou — Číselné řady označují hodnoty koeficientů při 95% a 99% pravděpodobnosti

Počet opakování ve zkupině / Počet zkupin	2	3	4	5	6	7	8	9	10
2	3,43	1,91	1,63	1,53	1,50	1,49	1,49	1,50	1,52
	7,92	3,14	2,47	2,24	2,14	2,10	2,08	2,09	2,09
3	2,37	1,44	1,25	1,19	1,18	1,17	1,17	1,18	,120
	4,42	2,14	1,74	1,60	1,55	1,53	,152	1,53	1,55
4	1,98	1,13	1,01	0,94	0,92	0,92	0,94	0,96	0,97
	2,96	1,57	1,33	1,24	1,21	1,21	1,21	1,22	1,23
5	1,40	0,94	0,85	0,81	0,80	0,80	0,81	0,82	0,84
	2,06	1,25	1,08	1,02	0,99	0,99	0,99	1,00	1,01
6	1,16	0,81	0,75	0,69	0,69	0,69	0,70	0,71	0,72
	1,69	1,04	0,94	0,86	0,85	0,85	0,85	0,85	0,86
7	1,00	0,70	0,63	0,61	0,61	0,61	0,62	0,63	0,63
	1,39	0,89	0,78	0,75	0,74	0,74	0,74	0,75	0,76
8	0,87	0,62	0,57	0,55	0,55	0,55	0,55	0,56	0,57
	1,20	0,78	0,69	0,66	0,65	0,65	0,66	0,66	0,67
9	0,78	0,56	0,51	0,50	0,49	0,50	0,51	0,51	0,52
	1,03	0,71	0,62	0,59	0,59	0,59	0,59	0,60	0,61
10	0,70	0,51	0,46	0,45	0,45	0,45	0,46	0,47	0,47
	0,91	0,62	0,57	0,54	0,53	0,54	0,54	0,55	0,55

uvádí *tabulka 2* hodnotu koeficientu $k = 1{,}18$, čili testovací kritérium.

$$r = k \,.\, \Delta = 1{,}18 \,.\, 26 = 30{,}68$$

Protože rozdíly mezi součty údajů hodnocení (631, 622, 644) jsou nižší než testovací kritérium, nemůžeme podle této metody považovat rozdíly v hodnocení za statisticky významné.

Kromě bodovacího systému je možno pro posouzení jakosti užít i způsobů dalších. Pro zjišťování rozdílů jakosti jsou zvláště vhodné tzv. párové zkoušky, metoda duo—trio a trojúhelníkové zkoušky. Při těchto způsobech hodnocení se nevyjadřuje jakost a její jednotlivé složky, ale od hodnotitele se požaduje vyjádření, který ze vzorků považuje v daném kritériu za lepší a musí vzorky několikrát od sebe odlišit. Pro jednoznačnost závěrů je nejsprávnější, když se oba porovnávané vzorky různí pouze v jediném faktoru, např. v surovině, technologii apod. Výhodou těchto metod je jednoznačnost odpovědí a poměrně snadné matematicko-statistické zpracování výsledků. V praxi samozřejmě nic nebrání zkombinování těchto postupů s bodováním.

Nejjednodušší z těchto metod jsou párové zkoušky. Ze dvou vzorků jeden označíme za kontrolní. Na volbě v zásadě nezáleží. Je však výhodnější vybrat za kontrolní vzorek ten, který považujeme za běžnější. Tento vzorek je hodnotiteli samostatně podán jako uváděcí. Úkolem hodnotitele je v následujících párech správně označit oba vzorky a stanovit, který je lepší. Počet párů vyplývá z povahy úkolu a z požadované jistoty. Minimální počet hodnocení při požadované pravděpodobnosti je možno vyčíst z *tabulky 3*. Při stanovení rozsahu všech hodnocení se však musí brát v úvahu i adaptace smyslových orgánů. Tato okolnost má velký význam u piva, při jehož hodnocení je smyslová únava mimořádně vysoká [2, 13].

Při větším počtu hodnocení (nad 20) se výsledky zpracovávají metodou kritického průměru. V jiné úpravě tutéž metodu uvádí jako „one teil test" *A. J. Byer* [14].

$$k_p = \frac{p_s - p}{\Sigma_p}$$

$$\Sigma_p = \sqrt{\frac{p\,(1-p)}{n}}\,,$$

kde k_p je kritický průměr;
p_s — procento správných odpovědí;
p — procento správných odpovědí teoreticky očekávaných;
Σ_p — střední chyba p;
n — počet všech odpovědí.

V případě párových zkoušek je teoreticky očeká-

Tabulka 3
Hodnocení párovými zkouškami

Počet párových zkoušek	Potřebný počet správných odpovědí při pravděpodobnosti		
	95 %	99 %	99,9 %
6	6	—	—
8	7	8	—
10	9	10	—
12	10	11	—
14	11	12	14
16	13	14	15
18	14	15	17
20	15	16	18
25	18	20	22
30	21	23	25
35	24	26	28
40	27	29	32
45	30	32	35
50	33	35	38
60	38	41	44
70	44	46	50
80	50	52	56
90	55	58	62
100	60	63	67
200	113	116	121
300	165	169	175
400	218	223	230
500	270	276	284

vané procento správných odpovědí 0,5, takže po úpravě vzorec zní

$$kp = \frac{(p_s - 0{,}5)\sqrt{n}}{0{,}5}$$

Pro k_p $(= u_p)$ najdeme v tabulkách normálního rozdělení odpovídající hodnotu pravděpodobnosti. Tedy např. při 70 % správných odpovědí a 25 hodnoceních

$$k_p = \frac{(0{,}70 - 0{,}50)\sqrt{25}}{0{,}50} = 2{,}00$$

což podle tabulek odpovídá 95% pravděpodobnosti. Opět vlastně zjišťujeme, s jakou pravděpodobností hodnota přesáhne danou mez a doplněk do sta nebo $(1 - p)$ představuje pravděpodobnost s níž hodnota v daném intervalu bude ležet.

Pro zpracování menšího počtu vzorků bylo by třeba používat poměrně složitějšího způsobu, vycházejícího z binomického rozdělení náhodné veličiny. *Tilgner* [1] však uvádí i pro tento případ počty správných odpovědí potřebných pro 95%, 99% a 99,9% spolehlivosti (*tabulka 3*).

Dalším vhodným způsobem pro zjišťování rozdílů v jakosti je tzv. dvojpárová metoda nazývaná též duo-trio. Je to v podstatě párová zkouška, při níž je počet párů omezen na dva. Nejprve je podán uváděcí vzorek a po něm následující dva páry vzorků. Výsledky se nejčasněji zpracovávají metodou X^2.

$$X^2 = \frac{(X_1 - X_2 - 1)^2}{n};$$

kde X_1 je počet správných odpovědí;
X_2 — počet nesprávných odpovědí;
n — celkový počet odpovědí.

Aby mohl být výsledek považován za spolehlivý z 95 % nebo 99 % či 99,9 % musí X^2 přesáhnout 2,71, 5,41, nebo 9,55. Tak při 13 správných odpovědí z 18

Tabulka 4
Hodnocení metodou duo — trio

Počet hodnocení a hodnotitelů	Potřebný počet správných odpovědí při pravděpodobnosti		
	95 %	99 %	99,9 %
5	4	5	5
6	4	5	6
7	5	6	7
8	5	6	7
9	5	6	8
10	6	7	8
11	6	7	9
12	7	8	9
13	7	8	9
14	7	9	10
15	8	9	10
20	9	11	12
25	11	13	14
30	13	14	16
40	16	18	20
50	18	20	23
60	20	23	26
70	22	26	30
80	25	29	33
90	28	32	36
100	30	35	39

Tabulka 5
Hodnocení trojúhelníkovými zkouškami

Počet trojúhelníkových zkoušek	Potřebný počet správných odpovědí při pravděpodobnosti		
	95 %	99 %	99,9 %
4	—	—	—
6	5	6	—
7	5	6	7
8	6	7	8
10	7	8	9
12	9	9	10
15	10	11	12
18	11	12	13
21	12	14	15
24	14	15	16
27	15	16	18
30	16	18	19
35	18	20	22
40	20	22	24
45	22	24	26
50	25	26	28
60	29	30	32
70	33	35	37
80	37	39	41
90	41	43	45
100	45	47	49

$$X^2 = \frac{(13 - 5 - 1)^2}{18} = 2{,}72$$

čili máme 95% jistotu, že výsledek je správný. Požadovaný počet správných odpovědí při dané pravděpodobnosti podle *Tilgnera* [1] je uveden v *tabulce 4.*

Poslední z běžných způsobů zjišťování rozdílů v jakosti je tzv. trojúhelníková metoda. Úkolem hodnotitele je vybrat ze 3 předložených vzorků, z nichž 2 jsou stejné, třetí nepárový. Je to úloha pro hodnotitele náročnější než předešlé, klade i větší nároky na dodržení správných podmínek, stačí však menší počet hodnocení pro získání statistiky spolehlivých výsledků.

Při počtu vzorků větším než 20 je možno hodnocení zpracovat metodou kritického průměru. Byla uvedena již ve zpracování párových zkoušek. Teoreticky očekávané procento správných odpovědí je 0,33, takže

$$k_p = \frac{(p_s - 0{,}33)\sqrt{n}}{0{,}47},$$

kde k_p je kritický průměr;
p_s — procento správných odpovědí;
n — celkový počet odpovědí.

Při 60 % správných odpovědí z 25 hodnocení

$$k_p = \frac{(0{,}60 - 0{,}33)\sqrt{25}}{0{,}47} = 2{,}87,$$

a to podle tabulek hodnot normálního rozdělení odpovídá jistotě 99,6 %.

Pro vyhodnocení nižšího počtu zkoušek než 20 udává *Tilgner* [1] opět přehled nutného počtu správných odpovědí v závislosti na celkovém počtu hodnocení a zvolené pravděpodobnosti (*tabulka 5*).

Závěr

Jak bylo již zdůrazněno v úvodu, účelem tohoto příspěvku je stručnou a přehlednou formou upozornit pracovníky pivovarsko-sladařského oboru na možnosti, které skýtá v organoleptice použití matematicko-statistických metod. Aby se usnadnila aplikace zásadních principů, byla pozornost zaměřena kromě několika všeobecných teoretických poznámek především na praktické zásady a použití. Proto také jsou ke článku připojeny i snadno použitelné tabulky uváděné v literatuře, z nichž bez velkých znalostí a dlouhých výpočtů lze okamžitě posoudit věrohodnost získaných podkladů. Většina z uváděných postupů je samozřejmě použitelná i při hodnocení jiných zkoušek než organoleptických. V tomto směru se doporučují některé z učebnic nebo příruček.

S rostoucími nároky na výrobu rostou neustále i nároky na její kontrolu a řízení. Tyto jsou tím účinnější, o co více se opírají o objektivně získané podklady a o co méně je používáno subjektivních odhadů. Při hodnocení potravinářských výrobků se ještě dlouho neobejdeme bez smyslového posouzení. Musíme proto provést všechna dostupná opatření, aby získané výsledky se co nejvíce blížily skutečnosti. Vedle správného smyslového posouzení to do značné míry může zabezpečit i matematicko-statistické zpracování výsledků.

Literatura

[1] Tilgner, J. D.: Organoleptická analýza potravin. SVTL, Bratislava 1961.
[2] Macher, L.: „Der Brauer und Mälzer", 4, 1966: 31—36.
[3] Byer, A. - Saletan, L.: „Wallerstein labor. communic.", XXV, 87, 1962: 309—323.
[4] Gray, P.: „Wallerstein labor. communic.", XXV, 87, 1962: 228—241.
[5] Robek, A.: Stanovení metodiky pro vyhodnocování přesnosti kontrolních metod organoleptických. Závěrečná zpráva Výzkumného ústavu pro mechanizaci a ekonomiku potravinářského průmyslu, Praha, č. 05.015 56, 1956.
[6] Štěpánek, V.: Matematická statistika v chemii. SNTL, Praha 1964.
[7] Šor, J. B.: Statistické metody analýzy a kontroly jakosti a spolehlivosti. SNTL, Praha 1965.
[8] Felix, N. - Bláha, K.: Matematicko-statistické metody v chemickém průmyslu. SNTL, Praha 1962.
[9] Reisenauer, R.: Metody matematické statistiky a jejich aplikace. SNTL, Praha 1965.
[10] Likař, O.: Statistické metody v laboratorní práci. SNTL, Praha 1957.
[11] Janko, J.: Teorie náhodných výběrů — odhady a některé testy významnosti. SNTL, Praha 1951. Učební texty vysokých škol.
[12] Tukey, J. W.: „American Society Quality Control Conf.", 5, 1951: 189—197.
[13] Laue, E. A. - Ishler, N. A. - Bullman, G. A.: „Food Technology", 8, 1954: 387—388.
[14] Byer, A. J.: „Food Technology", 18, 1964: 59—64.

Lektoroval ing. G. Klazar

Došlo do redakce 10. 10. 1966

МАТЕМАТИЧЕСКО-СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ОРГАНОЛЕПТИЧЕСКИХ ИСПЫТАНИЙ

Автор обращает внимание работников пивоваренно-солодильной промышленности на возможность использования современных математическо-статистических методов для анализа и обработки результатов разных органолептических испытаний. Приведены основы математического анализа и примеры его целесообразного применения.

MATHEMATISCH-STATISTISCHE AUSWERTUNG DER ERGEBNISSE ORGANOLEPTISCHER PROBEN

Der Autor macht in zusammenfassender und übersichtlicher Form die Mälzerei- und Brauerei-Fachleute auf die Möglichkeiten der Anwendung mathematisch-statistischer Methoden bei der Auswertung organoleptischer Proben aufmerksam. In dem Artikel werden die Grundmethoden sowie auch ihre Applikation angeführt.

APPLICATION OF MATHEMATICAL STATISTICS TO THE EVALUATION OF ORGANOLEPTIC TESTS

The article is addressed to the engineers employed in brewing industry and explains, how to evaluate the results of organoleptic tests by applying modern statistical and mathematical methods. Basic principles of the mathematical analysis are outlined and their application shown in typical examples.