

## XVI.

EN STATISTISK ANALYSE AV BAKEPRØVER.<sup>1)</sup>

Høsten 1935 fikk Universitetets Sosialøkonomiske Institutt anmodning fra direktør Osmundsen ved Kristiania Brødfabrik om å ta opp til undersøkelse visse statistiske spørsmål vedrørende brødproduksjonen. Instituttet erklærte seg villig. Det ble holdt en forberedende konferanse i september 1935 som nevnt i direktør Osmundsens artikkel «Godt og billig brød». Man bestemte seg til i første omgang å konsentrere seg om et produksjonsteknisk problem, nemlig å gjennomdrøfte statistisk en rekke *bakeprøver* som var utført på Kristiania Brødfabrik. Undersøkelsen ble bekostet av Brødfabrikken. Jeg ledet arbeidet; materiallbearbeidelsen og regnearbeidet ble utført av cand. oecon. Øivind Sundsbø. Arbeidet ble avsluttet i oktober 1936.

Den undersøkelsen som her er gjort, er på mange måter et skoleeksempel på hva en kan få ut av et tallmateriale når det underkastes en rasjonell bearbeidelse ved passende matematisk-statistiske metoder. En almenfattelig framstilling av gangen i undersøkelsen, og av de resultater en kom til, kan derfor kanskje være av interesse.

Materialet omfattet 46 bakeprøver fra Kristiania Brødfabrik. For hver prøve hadde vi følgende opplysninger. For *melets* vedkommende: proteininnhold, syregrad, vanninnhold og aske, alt regnet i prosent. For *deigens* vedkommende: deigutbyttet, gjæringsevnen, gjæringstoleransen og smidigheten. For *brødets* vedkommende: brødkvaliteten. Oppgavene over melets egenskaper og deigutbyttet var laboratoriemessige målinger, de andre oppgaver var skjønnsmessig fastsatte karakterer. Deigutbyttet er nærmest et mål for mengden av tilsatt vann. Vanntilsetningen i prosent er nemlig tilnærmelsesvis lik deigutbyttet minus 100. Resultatet av de 46 bakeprøver er gitt i tabellen. Her er målenhetene tilpasset slik at tallene får en bekvem form for det numeriske regnearbeide. Et høyt tall i rubrikk (9) betegner en høy brødkvalitet.

<sup>1)</sup> Publikasjon nr. 19 fra Universitetets Sosialøkonomiske Institutt.

1945

# 46 BAKEPRØVER FRA KRISTIANIA BRØDFABRIK

Prøve nr.	Melets				Deigs- utbyl- tet	Deigens			Brø- dets kvali- tet
	Pro- teininn- hold	Syre- grad	Vann- inn- hold	Aske inn- hold		Gjør- ings- evne	Gjør- ingstole- ranse	Smis- dighet	
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	
1	74	28	135	70	155	30	30	30	30
2	68	24	133	67	156	25	25	25	25
3	95	31	138	80	157	30	30	30	30
4	65	21	148	58	150	20	20	20	20
5	91	27	140	78	156	30	25	30	25
6	95	30	145	84	156	30	30	30	30
7	90	31	140	86	155	30	30	30	30
8	74	33	145	78	155	25	25	30	25
9	90	31	137	76	157	30	30	30	30
10	78	31	147	64	154	30	25	25	25
11	94	31	141	84	156	30	30	30	30
12	66	25	147	70	155	30	25	30	30
13	60	27	146	70	153	25	25	25	25
14	94	32	129	87	156	30	30	30	30
15	94	29	127	76	158	30	30	30	30
16	94	33	135	86	158	30	30	30	30
17	90	30	126	74	157	25	20	25	20
18	89	29	139	76	152	25	30	30	25
19	83	27	123	84	156	30	30	30	30
20	92	30	133	88	154	30	30	30	30
21	96	24	135	76	152	30	30	30	30
22	88	31	140	88	153	30	20	30	25
23	84	27	135	79	154	25	25	30	25
24	53	22	159	64	150	30	30	30	30
25	89	28	135	76	157	30	30	30	30
26	89	33	133	74	157	30	30	30	30
27	88	26	139	79	156	30	30	30	30
28	90	26	132	72	155	30	35	35	35
29	99	32	132	90	156	30	30	30	30
30	87	26	133	72	156	30	40	30	40
31	93	29	134	82	156	30	30	30	30
32	83	31	137	76	156	30	20	30	20
33	88	31	135	78	156	30	30	30	30
34	86	31	138	70	156	40	30	30	30
35	95	29	136	80	156	30	20	25	20
36	86	31	137	64	156	40	30	30	30
37	98	35	130	97	156	30	30	30	30
38	93	31	141	88	156	25	30	30	30
39	91	36	137	82	156	30	30	30	30
40	92	36	137	75	156	30	30	30	30
41	105	41	113	84	158	30	20	15	20
42	104	36	129	84	156	30	30	25	30
43	96	36	129	74	156	30	35	30	30
44	103	31	128	83	156	30	35	35	35
45	93	30	129	84	155	35	25	30	30
46	95	32	124	73	155	25	40	30	35

Kan tallene i denne tabellen lære oss noe om de faktorer som betinger et godt brød? Det første man bør gjøre når man stiller dette spørsmål, er å undersøke om de faktorene en her har opplysning om, er *uavhengige*, eller om noen av dem er så sterkt innbyrdes forbundne at de bare kan tas som uttrykk for et *enkelt* faktorkompleks.

Som et eksempel kan vi ta de to tingene deigutbyttet og vanntilsetningen, som ble nevnt foran. Det vilde være en forfeilet problemstilling å forsøke å finne ut hvorledes disse to tingene *hver for seg* kan bidra til en høy brødkvalitet. En må f. eks. forsøke å finne ut hvorledes brødkvaliteten vilde bli påvirket, om den ene av disse tingene ble holdt konstant, mens den annen ble endret. I virkeligheten er jo de to opplysningene bare forskjellige uttrykk for en og samme virkningsfaktor, nemlig mengden av tilsatt vann. Dette gjelder iallfall så lenge vi bare tenker på de *grove* trekk i billedet. Hvis analysen kunde bli drevet meget langt, ned til de fineste detaljer i årsakssammenhengen, da kunde det kanskje bli spørsmål om å betrakte vanntilsetningen og deigutbyttet som to *forskjellige* faktorer, f. eks. i den forstand at arbeidsprosessen under deigens framstilling kunde endres slik at for-dampningen blir enten større eller mindre, og at det dermed oppstår en variasjonsmulighet mellom deigutbyttet og vanntilsetningen. Men på analysens nåværende stadium, før en rekke andre og langt mere betydningsfulle faktorer er klarlagt, vilde en slik separatbehandling av vanntilsetningen og deigutbyttet bare lede oss på villspor uten den statistiske analyse.

Det er et spørsmål om ikke noe liknende kan gjelde for andre av de faktorene vi her har opplysning om. For å være på den sikre side bør man undersøke hva materialet selv kan si om dette. For å gjøre det bør man først jamføre variasjonene av faktorene to og to. Det kan f. eks. gjøres grafisk. En slik oversikt brakte for dagen at det er en nokså tydelig negativ sammenheng mellom deigutbyttet og melets vanninnhold, dvs. jo mindre melets vanninnhold er, desto større blir deigutbyttet, dvs. desto større prosent vann blir det tilsatt under framstillingen av deigen. Det er jo et nokså nærliggende forhold. Sammenhengen er imidlertid ikke så stramt at en uten videre må avvise melets vanninnhold som en ut-

sagnskraftig faktor ver siden av deigutbyttet. Men forsiktighet må utvises under den statistiske analyse.

En annen sammenheng som på denne måte kom til syne, var den mellom meleets proteininnhold og meleets vanninnhold. Mel med stort proteininnhold hadde gjennomgående lite vanninnhold. En tendens til negativ samvariasjon vil det selvsagt alltid være mellom to faktorer som begge er beregnet som prosentener innenfor en gitt masse (summen av alle prosentener er jo 100). Men i dette tilfelle var styrken i utslaget mere enn det som kan forklares på denne måte. En prosents oppgang i vanninnhold var nemlig gjennomsnittlig ledsaget av *mere* enn en prosents nedgang i proteininnhold (materialet viste gjennomsnittlig 1,44 ganger så sterk bedring i proteininnhold som i vanninnhold). Ingeniør Schulerud ved Kristiania Brødfabrik fant en slik sammenheng helt plausibel, idet f. eks. den argentinske rug, som har et høyt proteininnhold samtidig pleier å ha et lavt vanninnhold bl. a. p.g.a. dyrkestedets klima, mens f. ks. den tyske rug, som er proteinfattig, samtidig pleier å ha et stort vanninnhold. Dette er et moment til som gjør at en må være varsom med å trekke inn meleets vanninnhold som en selvstendig faktor under analysen.

Det ble også foretatt jamføringer mellom de forskjellige faktorene tre og tre; men det ble her ikke påvist noen nye avhengigheter som skulde kunne hindre oss i å betrakte faktorene som i det vesentlige uavhengige. Da kan det heller ikke eksistere slike avhengigheter i noen høyere komplekser av faktorene, f. eks. fire og fire osv. Den egentlige analyse av måten hvorpå *brødkvaliteten* avhenger av de forskjellige faktorene, kunde derfor begynne.

Det mest nærliggende er først å se på *deigens* egenskaper, og spørre hvilke egenskaper ved deigen det er som betinger det gode brød. Så kan en etterpå spørre om hvilke melegenskaper som bestemmer de forskjellige deigegenskaper.

Allerede en grafisk orientering om sammenhengen mellom brødkvaliteten og hver enkelt av deigegenskapene, en ad gangen, var nok til å vise at det var en temmelig stram sammenheng mellom brødkvaliteten og *deigens gjæringstoleranse*. En behøver i grunnen bare å la øyet gli nedover kolonnene

(7) og (9) i tabellen for å få et inntrykk av det. Det er også en viss sammenheng mellom brødkvaliteten og deigens *gjæringsevne*; men denne er ikke på langt nær så tydelig. I virkeligheten er sannsynligvis begge disse faktorene virksomme samtidig, slik at de utslag i brødkvaliteten som ikke kan forklares ved den ene av dem, kan forklares ved den annen. For å bringe rede i dette kan vi som en første tilnærming simpelthen oppfatte brødkvalitetskarakteren som sammensatt av to addender som hver for seg representerer virkningen av henholdsvis gjæringsevnen og gjæringstoleransen. Og videre kan vi for enkelhets skyld anta at hver av disse virkningene er en simpel proporsjonalitetsvirkning, dvs. framkommet ved at det tallet som representerer vedkommende faktor, multipliseres med et fast forholdstall. Dette forholdstall kommer da til å gi et uttrykk for vedkommende faktors *virkningsgrad*. Vet et statistisk beregningsarbeid, som det ikke er nødvendig å komme inn på i denne forbindelse, ble virkningsfaktorene for gjæringsevnen og gjæringstoleransen bestemt til henholdsvis 0,49 og 0,86. Det nivå som gjæringsevnen og gjæringstoleransens virkninger kommer i tillegg til, ble samtidig bestemt til — 10,4. Dvs. man stillet opp følgende formel hvoretter man skulde kunne *beregne* brødkvalitetskarakteren ut fra deigens gjæringsevne og gjæringstoleranse.

$$(1) \text{ beregnet brødkvalitet} = \begin{array}{l} 0,49 \times \text{gjæringsevne} \\ + 0,86 \times \text{gjæringstoleranse} \\ - 10,4 \end{array}$$

Som et eksempel kan vi ta prøve nr. 30, som ga den høyeste brødkvalitet. Her er gjæringsevnen = 30 og gjæringstoleransen = 40, altså  $0,49 \times 30 + 0,86 \times 40 - 10,4 = 38,7$ . Den beregnede brødkvalitet her er altså 38,7 mot den faktisk iakttatte kvalitet, som var 40. Som et annet eksempel kan vi ta prøve nr. 35, som er en av de dårligste. Her får vi  $0,49 \times 30 + 0,86 \times 20 - 10,4 = 21,5$ , mot den iakttatte brødkvalitet på 20. I betraktning av at brødkvaliteten bare er gitt i de store sprang 20, 25, 30 osv. er denne overensstemmelse mellom den beregnede brødkvalitet og den faktisk iakttatte meget god. Også for de andre prøvene vil en finne at overensstemmelsen er bra, unntatt for to av prøvene (nr. 34 og 36) hvor det formodentlig har gjort seg gjeldende spesielle forhold.

De faste forholdstall i formel (1) er bestemt ved en metode som tar sikte på å finne fram til det underliggende *typiske*, det som sannsynligvis har gyldighet uavhengig av de mer eller mindre tilfeldige feil og skjevheter som måtte forekomme i *dette* materiale.<sup>1)</sup> Hvis vi bare hadde stilt oss som oppgave å finne en formel som gjorde det mulig å beregne brødkvaliteten med størst mulig nøyaktighet ut fra gjæringsevnen og gjæringstoleransen i *dette* materiale, vilde vi ha funnet litt andre størrelser på de faste forholdstall, og overensstemmelsen mellom beregnet og faktisk iaktatt brødkvalitet vilde ha blitt enda litt bedre; men denne formelle økning av overensstemmelsen vilde ha hatt liten interesse.

Vi kan også beregne brødkvaliteten ut fra gjæringstoleransen og deigens smidighet; men resultatet blir stort sett ikke så godt som ved formel (1). Og om vi forsøker oss med en likning hvor alle de tre deigegenskaper gjæringsevne, gjæringstoleranse og smidighet forekommer samtidig, viser den statistiske analyse at bestemmelsen av de faste virkningsfaktorer blir temmelig upålitelig. Vi gjør derfor best i å bli stående ved formel (1), idet vi erindrer at gjæringstoleransen er den viktigste av de to faktorer i formelen. Ved eventuelle undersøkelser senere burde man forsøke å få bedre data, slik at man kunde få tatt hensyn til alle de tre deigegenskaper samtidig.

\*

Det neste skritt i behandlingen av det forelagte materiale ble å undersøke hvorledes meleets egenskaper og vanntilsetningen (målt ved deigutbyttet) virker inn på deigens gjæringsevne og gjæringstoleranse. Vi tar gjæringstoleransen først da den er den viktigste. En analyse av materialet viser at den uten sammenlikning viktigste faktor som har innflytelse på deigens gjæringstoleranse, er *melets proteininnhold*. Jo høyere proteininnhold, desto større gjæringstoleranse. Ingeniør Schulerud bekrefter at dete stemmer med de direkte erfaringer på fabrikk.

Melets askeinnhold viser seg også å ha en betydelig virkning på deigens gjæringstoleranse; men her foreligger et forhold som kan gi anledning til feilslutning hvis en ikke er

<sup>1)</sup> I den matematisk-statistiske terminologi: Vi har bestemt en strukturlikning i motsetning til en elementær regresjonslikning.

på vakt. Jmfører vi uten videre askeinnholdet og gjærings-toleransen, så viser materialet en positiv samvariasjon, dvs. jo større askeinnhold, desto større gjæringstoleranse. Denne sammenhengen gir imidlertid ikke uttrykk for noe selvstendig årsaksforhold, men bare for den omstendighet at det store proteininnholdet som skaper en stor gjæringstoleranse samtidig som regel er forbundet med et stort askeinnhold. Hvis vi får *isolert* virkningen av askeinnholdet, viser den seg å gå i motsatt retning. Når proteininnholdet holdes konstant, vil altså gjæringstoleransen bli desto større, jo *mindre* askeinnholdet er. Dette kan tydelig påvises i materialet. Det kan bringes til uttrykk ved at gjæringstoleransen uttrykkes i dens samtidige avhengighet av proteininnhold og askeinnhold ved en liknende formel som (1). Vi finner

$$(2) \text{ Beregnet gjærinstoleranse} = \begin{array}{r} 0,57 \times \text{proteininnhold} \\ - 0,76 \times \text{askeinnhold} \\ + 37,4 \end{array}$$

For den beste bakeprøven, nr. 30 får vi f. eks.

$0,57 \times 87 - 0,76 \times 72 + 37,4 = 32,27$  mot en faktisk iaktatt gjæringstoleranse på 40. Overensstemmelsen er ikke særlig god, og det er også andre prøver med nokså stor avvikelse mellom beregnet og iaktatt gjæringstoleranse; men tendensen i retning av negativ virkning av askeinnholdet er iallfall tydelig.

En nærmere analyse av materialet viser at det er en tredje faktor som her bør regnes med ved siden av de to faktorer i (2), nemlig meleets syregrad. Materialet viser at gjærings-toleransen er desto større, jo mindre syregraden er. I den rapporten om den statistiske undersøkelse som ble skrevet i 1936, antydet jeg at det for syregradens virkning formodentlig gjelder en optimumslov, slik at det er en viss midlere syregrad som er den gunstigste. Når det foreliggende materiale viste at gjæringstoleransen — og dermed brødkvaliteten — ble bedre ved en senkning av syregraden, så måtte det tolkes derhen at syregradens optimumspunkt i de foreliggende prøver i de fleste tilfelle var passert, dvs. i *de forelagte prøver var syregraden gjennomgående holdt for høy*. Denne konklusjon var for fabrikkene ikke så lite av en overraskelse, da den inntil da hadde gått ut fra at det bare gjaldt å sørge for å bruke mel hvis syregrad lå over et visst erfaringsmessig

fastlagt minimum, mens den hadde trodd at det var uten praktisk betydning om syregraden lå atskillig høyere enn minimum. Riktigheten av Instituttets beregninger på dette punkt er blitt bekreftet, og har fått sin forklaring ved de resultater den laboratoriemessige forskning er kommet til i årene etter 1936. Herom henvises til den interessante utredning som er tatt inn i direktør Osmundsens forannenvte artikkel.

La oss dernest se på hva det statistiske materiale kan si om de faktorer som betinger deigens *gjæringsevne*. Den synes å henge sammen med *deigutbyttet*, dvs. med mengden av tilsett vann; jo mere vann som tilsettes, desto bedre synes gjæringsevnen — og dermed brødkvaliteten — å bli. Slik ytrer sammenhengen seg når en jamfører gjæringsevnen og deigutbyttet uten videre, altså uten å ta hensyn til noen av de andre faktorene. Spørsmålet er om denne sammenhengen vil bli stående også når vi eliminerer virkningen av de andre faktorene, slik at vi får fram deigutbyttets (vanntilsetningens) virkning (*isolert* eller om virkningens retning vil slå om på liknende måte som askeinnholdets virkning slo om i dets forhold til gjæringstoleransen.

For å klargjøre dette må vi se på hvorledes gjæringsevnen avhenger av visse andre faktorer. Først og fremst viste det seg at gjæringsevnen ble gunstig påvirket av proteininnholdet. En likning som tar med både proteininnholdet og deigutbyttet som faktorer, viser seg imidlertid ikke å kunne brukes, da materialet ikke tillater noen utsagnskraftig bestemmelse av virkningskoeffisientene i en slik likning. Men hvis vi dessuten tar med melets askeinnhold, får vi en likning som iallfall kan vise *tendensen* i sammenhengen. Denne sammensatte likning, som skulde angi *nettovirkningen* av hver enkelt faktor, viser at gjæringsevnen — og dermed brødkvaliteten — blir desto høyere jo større deigutbyttet, proteininnholdet og vanninnholdet er, og jo mindre askeinnholdet er. Virkningsretningen for deigutbyttet har altså ikke slått om. Men askeinnholdets virkning slo om på liknende måte som ved (2).

Disse slutningene kan verifiseres ved at vi studerer *direkte* hvorledes brødkvaliteten avhenger av melets egenskaper og deigutbyttet. Jeg skal ikke gå i detalj, men bare nevne at



det ble bestemt en likning som rommer det vesentligste av hele årsakskomplekset, nemlig en likning hvorved brødkvaliteten beregnes ut fra de tre faktorer proteininnholdet, syregraden og deigutbyttet. I denne likning virker proteininnholdet og deigutbyttet positivt — som en måtte vente etter resultatene av den tidligere analyse via gjæringstoleransen og gjæringsevnen —, og syregraden virker negativt. Dvs. brødkvaliteten blir iflg. denne likning desto bedre, jo større proteininnholdet og deigutbyttet er og jo mindre syregraden er. Av den tidligere analyse vet vi at proteininnholdets virkning går både via gjæringstoleransen og gjæringsevnen (fortrinsvis via den første), mens syregradens vesentlig går gjennom gjæringstoleransen, og deigutbyttets (vanntilsetningens) vesentlig gjennom gjæringsevnen.

Både den direkte analyse av brødkvaliteten og den indirekte via deigens gjæringsevne viser altså en tydelig positiv virkning av deigutbyttet, dvs. jo større deigutbyttet er, (mens de andre faktorer er konstante) desto bedre blir brødkvaliteten. Ut fra produktivitetsteoriens generelle prinsipper skulde en vente at det for deigutbyttet (vanntilsetningen) vilde gjelde en optimumslov, dvs. slik at brødkvaliteten forbedres etter som vanntilsetningen til deigen økes opp til et visst punkt, for deretter å avta (alt mens de andre faktorer holdes konstante). En slik krumlinjet sammenheng kan selvsagt ikke påvises ved lineære likninger av den slags som er omtalt foran. Det ble derfor også brukt en annen og mere komplisert metode<sup>1)</sup> som er i stand til å påvise krumlinjetheten i sammenhengen hvis det er noen. Den ble brukt til å bestemme sammenhengen mellom brødkvaliteten og deigutbyttet når proteininnholdet og syregraden holdtes konstant lik de verdier som disse faktorer hadde gjennomsnittlig i det forelagte materiale. Resultatet ble ikke en kurve med noe tydelig toppunkt, men nærmest en serie med jevnt stigende punkter gjennom hele materialet. Dette resultat måtte tolkes derhen at bakeprøvene i det foreliggende materiale gjennomgående lå på den stigende kurvegren (før optimumspunktet), dvs. *det var gjennomgående brukt for liten vanntilsetning til deigen*. Etter at Instituttet hadde formulert denne slutning, fikk vi vite at dette

<sup>1)</sup> Den såkalte punktinterpolasjonsmetode.

stemte med de praktiske erfaringer ved fabrikken. Av hensyn til den maksimale behandling var vanntilsetningen ved disse prøvene holdt svært lav. Ingeniør Schulerud har også meddelt at han finner det helt plausibelt at den gunstige virkning av en økt vanntilsetning særlig skulde foregå gjennom en forbedring av deigens gjæringsevne slik som beregningene viste. Han mener det kan forklares på den måten at gjærcellene utvikler gass som må arbeide mot et visst trykk. Dette trykk blir mindre, jo løsere deigen er, dvs. jo større vanninnhold den har.

\*

Den utførte statistiske analyse bragte også for dagen en rekke detaljopplysninger av forskjellig slag, som det imidlertid vilde føre for langt å komme nærmere inn på her. Det som er anført, skulde være nok til å antyde at det her foreligger et stort og fruktbart felt hvor en sikkert vil kunne få mange verdifulle resultater ved et videre samarbeid mellom bakeri-, mølle- og korndyrkningseksperter på den ene side og matematiske statistikere på den annen. Jeg kan ikke annet enn å lykkønske Kristiania Brødfabrikk med det initiativ den her har tatt, og håpe at dette arbeid vil slå seg igjennom i videre kretser.