

# A Biometria na Seleccção

Eng. Agron. AGESILAU A. BITANCOURT  
do Instituto Biologico de Defesa  
Agricola Animal

No ultimo numero da Revista de Agricultura o professor Carlos Mendes expoz uns interessantes estudos de biometria, levantando a questão da verdadeira "significação das curvas de frequencia na representação dos phenomenos biologicos".

Tendo eu mesmo praticado intensivamente a biometria nos meus trabalhos de seleccção de plantas forrageiras quando dirigi a Estação Experimental de Agrostologia do Ministerio de Agricultura em Deodoro, não podia deixar de me interessar pelos estudos do professor Mendes. Como elle, tive na biometria fundas esperanças que mais tarde, no decorrer dos meus trabalhos, foram se transformando em duvidas. Devo dizer, entretanto, que as duvidas que me assaltaram não foram as mesmas que as do Dr. Mendes. Antes pelo contrario, procuro no presente artigo mostrar como, ao meu ver, estas ultimas podem ser levantadas. Mas antes de entrar propriamente na discussão dos resultados apresentados pelo professor da Escola Agricola é indispensavel examinar alguns pontos theoreticos da estatistica mathematica.

\* \* \*

A curva normal de frequencia apresenta-se em uma quantidade de problemas quer puramente theoreticos, quer praticos, em questões de mecanica, physica, sociologia, finanças, biologia, etc. Em theoria é ella a expressão da probabilidade ou da frequencia das diversas combinações possiveis entre diferentes proporções de dois factores reunidos em uma prova. Será por exemplo, no caso do problema da urna, a probabilidade de extrahir uma determinada proporção de bolas pretas e brancas em uma prova isto é, em uma extracção, de um numero muito grande destas mesmas bolas contidas em numero infinito e em determinadas proporções, no interior de uma urna. Sabemos

que se o numero de bolas extrahidas de uma só vez não for muito grande, 10 por exemplo, as probabilidades para as diversas combinações possíveis (neste exemplo, 11 combinações diferentes) são exprimidas pelos termos do desenvolvimento do binomio de Newton :

$$(p + q)^{10}$$

onde p e q exprimem as proporções respectivas das bolas pretas e brancas. Quando o numero de bolas extrahidas tende ao infinito, a expressão acima tende para uma expressão limite que é justamente a lei normal de probabilidade ou lei normal de frequencia.

A curva normal é continua. Podemos, para estudal-a, dividir-a em segmentos no interior dos quaes podemos considerar o seu valor como fixo. Substituíl-a-hemos assim por um polygono. Este, quanto menor forem os segmentos, mais se aproximará da curva theorica para finalmente confundir-se com esta ultima quando os segmentos tornam-se infinitamente pequenos. As tabellas biometricas nos fornecem os diversos valores da curva normal em funcção do desvio standard; podemos portanto fazer a divisão em polygonos tomando os segmentos da curva em correspondencia com segmentos das abscissas que sejam uma fracção ou um multiplo do desvio. E' o que fiz para segmentos das abscissas medindo duas vezes, uma vez e metade do desvio standard, obtendo as series seguintes que exprimem a altura dos polygonos, ou então a frequencia para mil variantes, ou ainda, dividindo por mil, a probabilidade para cada uma das combinações que estes segmentos podem exprimir :

Duas vezes						23	477	477	23					
Uma vez						12	22	135	341	341	135	22	12	
Metade	1	5	17	44	92	150	191	191	150	92	44	17	5	1

(Todos os numeros foram arredondados para tornal-os inteiros).

Para comprehender o que segue, não devemos esquecer que qualquer destas series exprime a lei normal, embora ellas apresentem distribuições bem diversas das mil variantes.

Sobre estes principios theoreticos assenta-se a estatistica ma-

thematica e como a biometria é a aplicação da estatística theorica aos problemas biologicos, comprehende-se que o problema da urna está na base da biometria.

A razão pela qual a estatística mathematica tem applicação na biologia está na observação de que innumerous caracteres chamados fluctuantes, isto é, caracteres não fixos, caracteres que podem assumir diversos valores numa mesma população e até num mesmo individuo, — nos reinos animal e vegetal, variam justamente de accordo com a lei normal de frequencia. Trata-se ahi de um facto concreto, resultado de observações muitas vezes repetidas pelos scientistas os mais notaveis e que não é mais passivel de contestação. Para explicar este phenomeno admite-se que innumerous factores, alguns favoraveis e outros desfavoraveis, determinam o valor final do caracter estudado. Os diversos valores encontrados se explicam pelo facto dos differentes factores agirem á maneira das diversas bolas presentes nas extracções da urna, combinando-se em differentes proporções para dar nos seus effeitos a curva normal da frequencia.

Vejamos como se estuda biometricamente um determinado caracter. Supponhamos que se trate de um caracter susceptivel de variação continua, como por exemplo a altura ou o peso de uma planta ou de um animal, ou de uma parte dos mesmos. Desde já apresenta-se a questão do valor ou modulo das classes que vamos estabelecer para repartir os individuos.

Adoptando determinado modulo, ou outro de valor duplo, triplo, etc., fazemos exactamente a mesma coisa do que dividindo a curva normal em polygonos, como acima, com uma unica differença, é que neste ultimo caso adoptamos como unidade das obscissas o desvio standard, ao passo que agora adoptamos grammas, centimetros, etc. Em ambos os casos temos ampla liberdade para escolher o modulo que queremos, porem quanto menor o modulo, mais nos approximamos da curva theorica, sem que sobretudo, (qualquer que seja o modulo adoptado), as series obtidas deixem de exprimir a curva normal, á semelhança das series que acima citei. Isto bem entendido se o caracter estudado segue realmente a lei normal.

Mas uma vez adoptado determinado modulo, — o que fi-

ca, é bom repetir, inteiramente ao nosso criterio, de accordo, em theoria, com o grau de precisão que desejamos, e em pratica de accordo com a facilidade e commodidade da medição e do tempo de que dispomos, — uma vez adoptado o modulo, somos obrigados a fazer um numero de medições que esteja em relação com o numero de classes que a escolha do modulo determinou. E' indispensavel que o numero de individuos, pelo menos nas classes onde a frequencia é maior, seja sufficientemente grande para que as leis de probabilidades se verifiquem realmente, sem o que o resultado obtido será perfectamente illusorio.

Para termos uma ideia da quantidade de individuos que devem entrar, em theoria, nas diversas classes, de accordo com o valor do modulo, basta considerar o exemplo das tres series que dei acima para modulos equivalentes a duas vezes, uma vez e metade do desvio standard. Ao passo que na primeira serie havia 477 individuos nas classes maiores, ha somente 341 e 191 respectivamente nas duas outras. Pode-se admittir que a precisão é inversamente proporcional ao erro e portanto directamente proporcional á raiz quadrada do numero de individuos medidos. E' pois indispensavel, para obter a mesma precisão medir um numero de individuos muito maior na ultima serie do que na primeira.

A experiencia mostra que para obter polygonos que representem bem a população donde foi tirada a amostra, é indispensavel, quando estes polygonos seguem a lei normal, incluir o numero seguinte de individuos medidos, de accordo com o numero de classes em que se encontram 99 % destes ultimos:

Numero das classes que contém os 99 % da amostra	Numero total dos individuos da amostra
5	100
7	200
10	400
15	900
20	1.600

Não ha em geral grande interesse em procurar ter mais de 20 classes porque não somente fica sendo muito alto o nu-

mero de variaveis necessarias como tambem porque 20 classes determinam um polygono muito approximado da lei normal, bem sufficiente para apurar a similitude. No caso de uma população heterogenea, entretanto, a multiplicação das classes permite apurar mais facilmente os diversos modos existentes, mas então é preciso não esquecer que o numero de individuos necessarios augmenta como o quadrado do numero das classes.

Os numeros que dou acima são numeros que garantem uma curva satisfactoria, mas bem entendido, ha casos em que numeros menores permitem obter um polygono sufficientemente regular.

Eu disse que os casos em que os caracteres fluctuantes obedecem á lei normal são muito numerosos em biologia. E' indispensavel saber reconhecer quando uma curva ou um polygono de frequencia obtidos empiricamente pela medição de determinado caracter approximam-se ou não da curva theorica. Procede-se para isto ao "ajustamento" da curva empirica, operação fundamental da biometria de que muitos livros de genetica que tratam da biometria muito por alto são completamente omissos. E' melhor não fazer biometria do que fazel-a por alto.

Para não alongar por demais o presente artigo não farei aqui a exposição de como se realiza este ajustamento, mesmo pelo processo elementar geralmente usado em biometria. Lembrarei tão somente que dando as tabellas a curva theorica em funcção do desvio standard, é preciso antes de tudo calcular este desvio e estabelecer o valor das classes tomando-o por unidade. E foi este o processo que empreguei no estudo das curvas do professor Mendes.

\*  
\* \*

Vejamos agora estas curvas. Começarei pela mais caracteristica que se refere ao Arroz Douradão. O prof. Mendes nos diz que praticou a selecção intensiva de uma variedade já previamente seleccionada, durante quatro annos consecutivos. Em uma planta de fertilização directa como o arroz deve-se presumir que a população tenha alcançado um estado de homozygose absoluta, tratando-se pois de uma linha pura em que a lei normal deve se verificar facilmente.

O Dr. Mendes estabeleceu o graphico III do seu trabalho e pergunta: "Revelará esse graphico uma linhagem pura, ainda que possuindo quatro vertices lateraes, ou porque elles existem esse facto é bastante para não se poder asseverar a existencia dessa linhagem?"

Ha evidentemente um equivoco nessa pergunta. A curva normal não pode nunca revelar a primeira vista uma linhagem pura, como veremos adeante. Ella só nos revela uma coisa, é a maneira por que varia o character. Acontece que nas linhas puras, em geral esta variação segue a lei normal, mas ha casos em que esta tambem se verifica sem que a linha seja pura e a biometria é incapaz de nos dizer no qual dos dois casos está o character estudado.

Feita esta restricção que será, aliás, discutida mais amplamente adeante, verificamos em primeiro lugar que para o numero pequeno de 393 variantes é completamente impossivel obter uma curva satisfactoria com cerca de 30 classes. Seriam necessarias cerca de 3.500 medições para obter um resultado plenamente satisfactorio. Deante disso, avaliando pela curva <sup>(1)</sup> as diversas frequencias obtidas pelo professor Mendes, empreguei um modulo tres vezes maior, sommando as frequencias de tres em tres classes.

Consegui assim reduzir a 10 o numero de classes, ficando portanto dentro das prescripções acima ditadas, porquanto tinha cerca de 400 individuos para nellas distribuir.

Encontrei a serie seguinte, e calculando a media e o desvio standard fiz o ajustamento do polygono obtido á curva normal de frequencia cujas ordenadas correspondentes serão inscriptas por baixo das respectivas frequencias: (Fig. 1)

Classes	1	2	3	4	5	6	7	8	9	10	( <sup>2</sup> )
Frequencia observada:	7	11	29	59	87	89	68	30	11	2	
Frequencia calculada:	3	11	31	63	88	88	63	31	11	3	

(1) E' possivel que nestas avaliações sobre os diagrammas do artigo do professor Mendes tenha se introduzido algum erro de pouca importancia, pela difficuldade de leitura. Isto explicará alguma differença que por ventura exista entre as minhas series e as do dr. Mendes. Estas differenças, de qualquer forma não podem ser de molde a alterar os resultados que apresento.

(2) As classes, nesta serie e nas seguintes, estão designadas por um simples numero de ordem. O exacto valor dessas classes pouco importa para o presente estudo.

Como se vê a concordancia é notavel. De facto nunca encontrei nos meus estudos e mesmo na litteratura um caso tão perfeito de ajustamento. A figura 1 representa a lei normal de frequencia trazendo na forma dos pequenos circulos os valores encontrados pelo Dr. Mendes.

Quanto aos vertices a que se refere o professor Mendes vê-se no decorrer do artigo que elle attribue a esses accidentes de suas curvas um valor que que absolutamente não têm. Considera-os como modos eventuaes, distinctos. Na realidade só se póde affirmar que uma curva é multimodal quando os di-

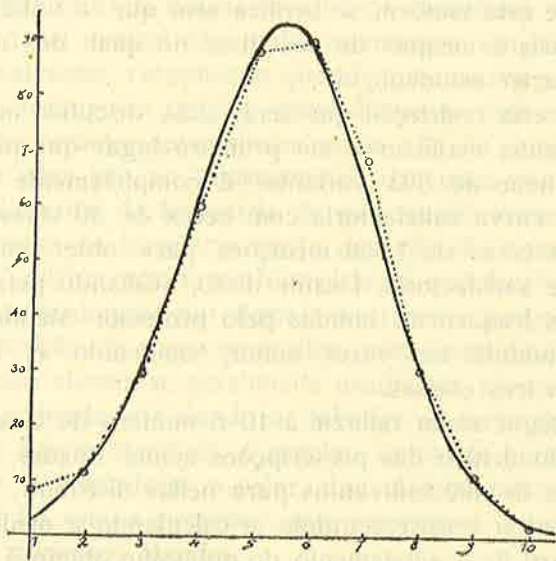


Fig. 1 — Curva de frequencia do arroz Douradão comparada com a curva normal. Cada unidade das ordenadas corresponde a uma variante

versos modos são reproduzidos com perfeição em diversas curvas obtidas numa população com diversas amostras (e não com a somma de diversas amostras como o fez o professor Carlos Mendes, introduzindo dest'arte uma correlação involuntaria entre as diversas curvas).

Procurando, por exemplo, continuar a selecção da linha 52 de capim de Rhodes (*Chloris gayana* Kunth) que eu obtive na Estação Experimental de Agrostologia e que eu suspeitava ser pura, escolhi tres pés que designei E. E. A. 121, E. E. A.

122 e E. E. A. 123, colhendo suas sementes que me forneceram tres canteiros visinhos numa mesma parcella de cultura. Realizei a biometria do numero de espigas por inflorescencia em cada um dos canteiros e obtive para o 121 e o 122 as duas curvas da figura 2 que foram traçadas na mesma escala, transpondo, tão somente as ordenadas referente á 122, 15 unidades acima. O parallelismo é evidente. Ha um modo muito nitido e outro muito menor que as duas curvas reproduzem perfeita-

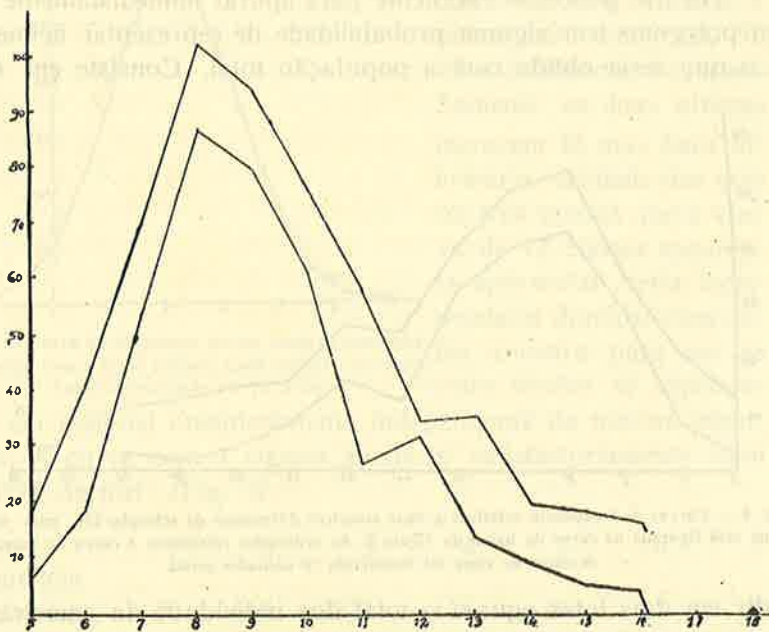


Fig. 2 — Curva de frequência do numero de espigas por inflorescencia nas selecções E. E. A. 121 (em baixo) e 122. As ordenadas referem-se a numero de variantes da selecção 121. Os numeros relativos á selecção 122 foram transferidos 15 unidades acima

mente. Estas curvas multimodales mostram que as duas populações 121 e 122 são identicas, pois conforme mostrarei adiante o numero de espigas por inflorescencia no capim de Rhodes obedece a uma fluctuação multimodal. A curva obtida com a linha 123, embora muito parecida com as duas primeiras apresenta algumas diferenças que talvez possam ser attribuidas a ligeiras diferenças de condições de meio. De facto os tres canteiros foram plantados na ordem respectiva e é possi-



vel que o 123 tenha encontrado condições ligeiramente diversas na extremidade da parcella. O que teria sido realmente extraordinario seria encontrar curvas paralelas para o 121 e o 123 e ligeiras diferenças para o 122 situado entre os dois outros. Das trez curvas obtidas é licito concluir que não foi possível extrahir linhas diversas da linha 52 a que pertencem genotypicamente as linhas 121, 2, 3, confirmando portanto a presumpção de que a 52 já era uma linha pura.

Ha um processo excellente para apurar immediatamente se um polygono tem alguma probabilidade de representar fielmente o que seria obtido com a população total. Consiste em di-



Fig. 3 — Curvas de frequência relativas a duas amostras diferentes da selecção 121, cuja somma está figurada na curva de baixo da figura 2. As ordenadas referem-se á curva de baixo. A curva de cima foi transferida 10 unidades acima

vidir em dois lotes eguaes o total dos individuos da amostra e fazer separadamente a biometria de cada um dos lotes. Se as curvas obtidas não são sufficientemente paralelas toma-se um lote supplementar até conseguir a semelhança desejada. Por exemplo, a biometria referente ao 121 deu as duas curvas da figura 3, para dois lotes de 200 individuos. Como se vê o parallelismo justificou a sufficiencia de 400 individuos embora as prescrições que acima apontei pedissem um numero maior. Nota-se ainda que o modo supplementar á direita é perfectamente significativo, confirmado que elle é nas duas curvas.

※

\* \*

Passemos para outro caso, o do clone de mandioca. Evi-

dentemente, excluindo a hypothese pouco provavel do apparecimento de alguma mutação vegetativa, a população que servio ao estudo do prof. Mendes pode ser considerada como composta de um unico individuo sob o ponto de vista genetico. Com o numero de 693 individuos vê-se que na base da tabella acima não se deveria passar de umas treze classes para conseguir um polygono satisfactorio. O prof. Mendes, entretanto,

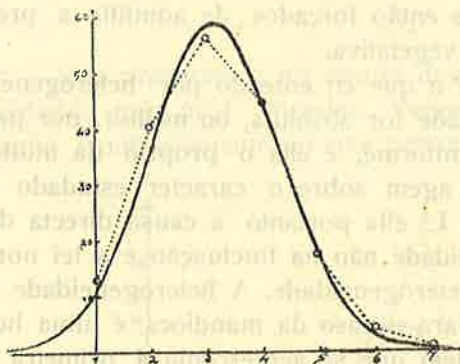


Fig. 4 — Curva de frequência de um clone de mandioca, comparada com a curva normal. Cada unidade das ordenadas corresponde a 4 variantes

estabeleceu curvas com cerca de cem classes e outras com umas 50, 20, 12 e finalmente 7 classes. Somente as duas ultimas merecem fé, mas para afirmar a realidade dos dois ou tres modos que a curva de 12 classes apparenta apresentar seria indispensavel duplicar com outra amostra para ver se estes modos se reproduzem

em material completamente independente da mesma população. A curva com 7 classes ajusta-se satisfactoriamente com a curva normal : (Fig. 4)

Classes	1	2	3	4	5	6	7
Frequencia observada	50	160	223	178	69	17	1
Frequencia calculada	44	146	232	183	70	13	1

No caso não ha motivo para decepção. A biometria, se de facto ha realmente dois ou tres modos em vez de um como se podia esperar, teve o alto merito de nos chamar a atenção sobre os mesmos, o que provavelmente nenhum outro methodo de pesquisa teria revelado. Mas a biometria nos ensina somente que ha heterogeneidade na população em estudo. Ella não nos indica qual a natureza desta heterogeneidade. Isto o biologista, de posse desta informação, deve procurar por outros methodos. No caso presente, caso ella seja comprovada pelo me-

thodo que indiquei, isto é, pela duplicação das amostras, ella pode ser attribuida a uma heterogeneidade notavel do meio em que foi feita a cultura: solo, exposição, etc. ou então do proprio material empregado e neste caso ainda teremos que distidguir a heterogeneidade que chamarei de morphologica como a de que fallarei adeante relativamente ás inflorescencias do capim de Rhodes e a heterogeneidade genetica, sendo que no caso de um clone seremos então forçados de admittir a presença de alguma mutação vegetativa.

E preciso definir aqui o que eu entendo por heterogeneidade. Se esta heterogeneidade for absoluta, ou melhor, por paradoxal que isto pareça, uniforme, é ella o proprio da multiplicitade de factores que agem sobre o character estudado e com que ella se confunde. E' ella portanto a causa directa da fluctuação. Sem heterogeneidade não ha fluctuação e a lei normal é a expressão desta heterogeneidade. A heterogeneidade a que eu me referia acima para o caso da mandioca é uma heterogeneidade de outra ordem que se acrescenta á primeira e que as curvas multimodaes revelam. Comparando com o que se vê na theoria dos erros, podemos dizer que esta heterogeneidade de segunda ordem é para a primeira o que o erro systematico é para o erro fortuito.

Vamos concretizar: o mandiocal do prof. Mendes tem duas zonas que são de meio differente, as raizes produzidas em cada uma das duas zonas consideradas separadamente são submettidas a uma infinidade de factores alguns favoraveis e outros desfavoraveis que constituem a *heterogeneidade fortuita*. Estas raizes fornecerão então uma fluctuação normal, isto, porem, em torno de modos differentes. Reunida a totalidade da colheita teremos forçosamente uma curva bimodal que revela uma *heterogeneidade systematica* <sup>(1)</sup>, no caso presente devida a existencia das duas zonas differentes.

Repetimos, a biometria nos revela portanto a heterogeneidade systematica, separando-a da heterogeneidade fortuita. Isto

(1) Está longe do meu pensamento procurar introduzir na linguagem biometrica as duas expressões de "heterogeneidade fortuita" e "heterogeneidade systematica", nem tão pouco considero-as bem acliadas. Emprego-as na falta de outras, para a commodidade de minhs exposição e porque lembram os erros systematicos e fortuitos da theoria dos erros que representam quantidades equivalentes.

já é uma grande coisa. O que ella não nos indica em absoluto é a natureza desta heterogeneidade. O mal de muitos que tentaram se servir da biometria tem sido de exigir desta muito mais do que ella pode realmente dar. De outro lado é forçoso convir que isto restringe muito o campo da biometria e concordo em que em muitos casos o seu auxilio é completamente illusorio!

\*  
\* \*

Mas prosigamos no estudo dos interessantes casos apresentados pelo prof. Mendes. Vejamos as bananas do mesmo cacho. Contrariamente ao que pensa muita gente, e sem appe-

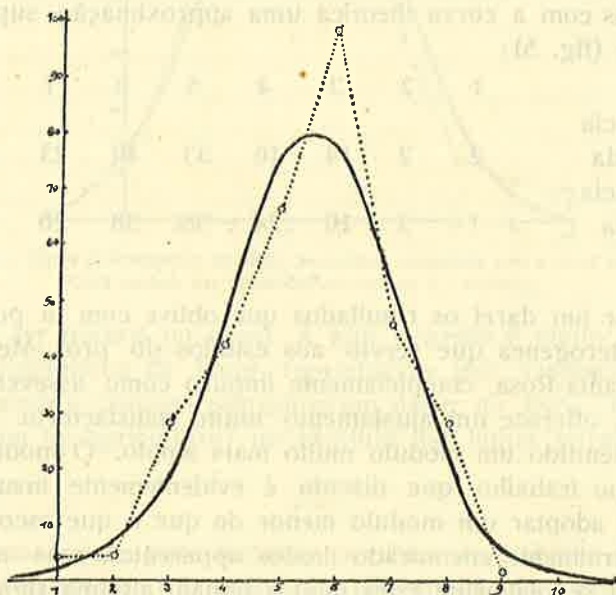


Fig. 5 — Curva de frequência das bananas de um mesmo cacho comparada com a curva normal. Uma variante é representada por duas unidades das ordenadas

lar para combinações genotypicas deseguaes como lembra o prof. Mendes é muito mais difficil obter uma curva regular com os elementos fornecidos por um unico individuo do que com um clone proveniente deste individuo. Em primeiro lugar é difficil conseguir um numero sufficiente de variaveis, em segundo lugar, pergunto se não é mais facil conseguir uniformidade

colhendo numa area restricta bem uniforme no solo e na exposiçãõ, as bananas da face exposta ao sol das pencas centraes dos cachos de diversas bananeiras pertencentes a um mesmo clone, do que colhendo todas as bananas de um mesmo cacho, tanto as que estãõ na parte de cima exposta ao sol, do que as que estãõ por baixo; as que estãõ na base do cacho (formadas as primeiras, livres de contacto com as outras), as do centro, formadas em seguida, ou as da extremidades, formadas por ultimo, prejudicadas no seu crescimento pela visinhança das outras desenvolvidas anteriormente. Por isto não nos devemos admirar se com um numero realmente insufficiente de variantes, que só permitem a distribuiçãõ em 6 classes, não conseguimos com a curva theorica uma approximaçãõ superior á seguinte (fig. 5):

Classe	1	2	3	4	5	6	7	8	9
Frequencia observada	2	2	14	16	33	49	23	14	1
Frequencia calculada	1	3	10	24	38	38	26	11	3

\*  
\*\*

Por fim darei os resultados que obtive com a populaçãõ mais heterogenea que servio aos estudos do prof. Mendes. O milho Santa-Rosa, completamente impuro como assevera o Dr. Mendes, offerece um ajustamento muito satisfactorio. Adoptei bem entendido um modulo muito mais amplo. O modulo adoptado no trabalho que discuto é evidentemente insufficiente. Poderia adoptar um modulo menor do que o que escolhi e teria naturalmente encontrado modos apparentes, mas nada me indicaria se realmente estes modos tinham alguma significaçãõ. Eis os resultados (fig. 6)

Classe	1	2	3	4	5	6	7	8	9
Frequencia observada	49	178	364	402	334	165	59	36	29
Frequencia calculada	60	163	307	395	346	208	87	24	4

Mas entãõ surge uma objecçãõ. Como é que uma populaçãõ heterogenea pode fornecer uma curva normal? Volto á

definição do que se deve entender por heterogeneidade. A biometria nos permite differenciar a heterogeneidade systematica da heterogeneidade fortuita. No caso de duas ou tres linhas puras misturadas numa população ha heterogeneidade systematica. No caso em que ha um numero muito grande de linhas misturadas uniformemente ha heterogeneidade fortuita e a cur-

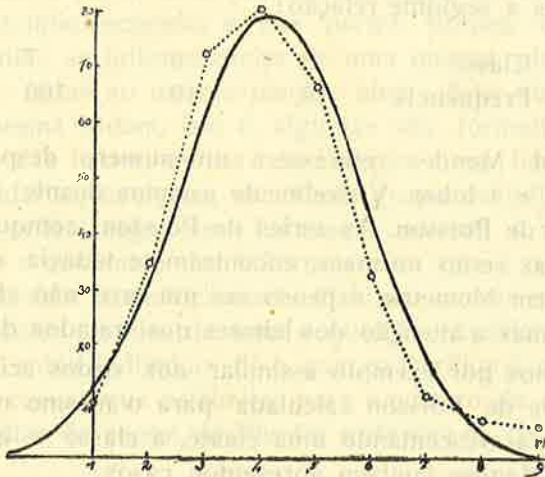


Fig. 6 — Curva de frequência do Milho Santa-Rosa comparada com a curva normal.  
Cada unidade das ordenadas corresponde a 5 variantes

va pode ser normal ou quasi. A este respeito é muito elucidativo o diagramma de Lang, reproduzido por Goldschmidt e que Babcock e Clausen reproduziram deste no livro "Genetics in Relation to Agriculture" no capítulo das linhas puras.

\*  
\*\*

Vamos agora para terminar a critica do trabalho do prof. Mendes considerar o exemplo por elle citado da Vassourinha. Agora não se trata mais de um caracter susceptivel de variação continua e sim do numero de lobos das folhas. Para explicar os casos, encontrados em biologia, em que caracteres não variaveis de um modo continuo e sim susceptiveis de tomar um certo numero de valores fixos, — que são geralmente as classes escolhidas para o estudo do caracter, — fornecem polygonos que se ajustam á curva normal, é forçoso admittir que os

factores que concorrem para fixar o caracter nas diversas classes, variam de um modo continuo embora o efeito apparente desta variação seja caracteres fixos. Admitte-se então que é necessario aos factores alcançar um certo valor para determinar a passagem de uma classe á classe immediatamente acima.

Na primeira medição foi encontrado para as folhas de 5, 6, e 7 lobos a seguinte relação :

Classe	5	6	7
Frequencia	10	10	100

O prof. Mendes refere-se a um numero desprezível de folhas de 3 e 4 lobos. Visivelmente estamos deante de um caso de serie de Poisson. As series de Poisson, comquanto mais raras que as series normaes, encontram-se todavia com certa frequencia em biometria e penso ser um erro não chamar sobre as mesmas a attenção dos leitores nos tratados de genetica.

Podemos por exemplo assimilar aos dados acima a seguinte serie de Poisson calculada para o mesmo numero de variaveis, e accrescentando uma classe, a classe 4 que segundo o prof. Mendes tambem apresentou casos :

Classes	4	5	6	7
Frequencia calculada	0,12	2	20	98

Caso uma duplicação confirmasse a differença entre estes numeros com os numeros observados, seria preciso admitir aqui uma influencia perturbadora, uma heterogeneidade systematica que veio alterar a previsão. Esta heterogeneidade está talvez no phenomeno que precisamos invocar para as series encontradas depois do florescimento :

Classe	5	6	7
Frequencia	35	10	100

Vemos ahi accentuada a heterogeneidade já revelada pela primeira serie quando comparada com a serie theorica de Poisson. E a mesma coisa, mais accentuada ainda na Vassourinha grande :

Classes	3	6	7
Frequencia antes do florescimento	25	20	100
Frequencia depois do florescimento	100	10	100

Trata-se ahi com certeza de curvas multimodas do tipo Poisson. E aqui preciso dizer algumas palavras das curvas multimodas encontradas numa mesma planta, em seus clones, ou numa linha pura. Estas curvas são frequentemente encontradas nas inflorescencias e nas partes floraes. Explicam-se perfeitamente: as inflorescencias de uma mesma planta não se constituem todas ao mesmo tempo; além disto ellas não têm todas a mesma ordem, isto é, algumas são formadas nas extremidades das hastes, outras lateralmente e em hastes primarias, secundarias, terciarias, etc. Cada grupo de inflorescencias de mesma ordem segue uma fluctuação normal, o mesmo não se dá com o conjuncto das inflorescencias que é o que geralmente se estuda biometricamente. Dahi as curvas multimodas desconcertantes obtidas de uma população considerada pura e até de um unico individuo. E' o que eu verifiquei em todas as innumeradas curvas que estabeleci para o numero de espigas nas inflorescencias do capim de Rhodes e de que eu dei um exemplo acima.

Não ha pois motivo para a admiração do prof. Carlos Mendes deante do resultado obtido com as vassourinhas. Acrescentarei que uma linha recta pode perfeitamente ser a expressão de uma curva normal.

Supponhamos uma população vegetal composta de plantas cujas flores possuem petalas nas seguintes proporções:

Numero de petalas	5	6	7
Proporção das flôres	25/2.000	1950/2.000	25/2.000

Esta serie corresponde á lei normal de probabilidade, tomando como modulo 4 vezes o desvio standard.

Para estudar o numero de petalas das flores desta população devemos empregar uma amostra, cem flores por exemplo. O calculo de probabilidades nos ensina que ha pouco menos de 6 probabilidades em mil <sup>(1)</sup> de tirarmos a combinação:

(1) A probabilidade de tirarmos a combinação

2

95

3

que melhor representaria a verdadeira composição da população em 100 variantes é tão somente 10 vezes maior ou pouco menos de 6 probabilidades em cem.



Numero de petalas	5	6	7
Frequencia	5	95	0

Dando-se esta eventualidade, teremos no lugar da curva ou melhor, do polygono esperado, uma recta como achou o prof. Mendes. Seria entretanto um resultado perfeitamente ad-sive, embora excepcional, e não haveria motivo para admiração. A hypothese de tirar a sorte grande numa loteria é infinitamente mais aleatoria, e entretanto é o que espera o jogador todas as vezes que compra um bilhete!

\*  
\*\*

E' necessario agora pôr uma ordem na somma de factos que estão acima expostos e comentados, e procurar responder á pergunta do prof. Carlos Mendes:

“QUAL A SIGNIFICAÇÃO DAS CURVAS DE FREQUENCIA NA REPRESENTAÇÃO DOS PHENOMENOS BIOLOGICOS”?

Conforme já disse acima a biometria só nos permite saber uma coisa: a população estudada apresenta ou não heterogeneidade systematica. No primeiro caso ella fornece uma curva irregular, geralmente multimodal (1). No segundo caso ella segue a curva normal de frequencia.

A heterogeneidade pode ser extrinseca, isto é ser dependente do meio como por exemplo a heterogeneidade que provem da presença de solos differentes no campo da cultura, da exposição differente das diversas plantas do campo ou das diversas partes do vegetal como no caso das bananas de um mesmo cacho, etc., ou então intrinseca e depender do proprio vegetal. E' então necessario distinguir a que provem da organização morphologica do vegetal como a que vimos para o capim de Rhodes da que é causada pela mistura de linhas geneticamente diversas. E' esta heterogeneidade que finalmente o geneticista procura apurar por meio da biometria, mas — não

(1) Seria aqui necessario dizer algumas palavras sobre as curvas leptocurticas e platycurticas e a possibilidade de se encontrar curvas unimodales em populações misturadas, porem com seus diversos modos reunidos numa só classe, mas não quero estender por demais o presente artigo.

é demais insistir sobre este facto, — a biometria é incompletamente capaz de fazel-o. Ella indica a heterogeneidade mas não a categoria a que ella pertence, das tres acima indicadas.

No caso em que o character estudado segue a lei normal a biometria ainda é muitas vezes incapaz de nos revelar, á primeira vista, se se trata de uma linha pura ou de uma mistura homogenea de um grande numero de linhas como vimos para o milho Santa Rosa.

Podemos resumir estes factos todos no quadro seguinte :

Heterogeneidade systematica (curva irregular)	}	Heterogeneidade mesologica Heterogeneidade morphologica Heterogeneidade genetica
Heterogeneidade fortuita (curva normal)	{	Linha pura Mistura homogenea de linhas

\*  
\* \*

Quasi ao terminar o seu artigo o prof. Mendes diz o seguinte:

“Ha ainda a considerar nestas experiencias um outro factor de grande importancia, o *modulo* a ser adoptado para avaliar as grandezas”.

Penso ter mostrado que nestas experiencias é quasi unicamente o modulo que deveria ter sido considerado. Quasi sempre, com certeza para obter uma precisão maior, completamente inutilizada pelo pequeno numero relativo de variantes empregado, este modulo foi pequeno demais, fornecendo curvas irregularissimas onde difficilmente era possivel chegar a alguma conclusão relativamente aos caracteres da população estudada. Vimos então como alterando este modulo para valores que estavam mais em relação com a importancia da amostra consegui boas approximações com a curva theorica, tirando dos dados apresentados pelo prof. Mendes tudo o que razoavelmente se pode exigir da biometria. Terminando direi que a biometria é um meio e não um fim. Muitos acharão que as minhas conclusões são magras e cheias de restricções nos casos estudados. E' que a biometria deve acompanhar a selecção e somente uma serie de curvas obtidas em diversas epocas desta selecção permitem ajuizar da natureza da população com que se trabalha, acompanhando o progresso da selecção na transformação desta população. De uma só curva não penso que um biometrista consciencioso se aventure a tirar muito mais conclusões de que eu tirei, das curvas isoladas do professor Carlos Mendes.