PEER REVIEWED

Sampling error in measurement of dirt in pulp and paper

JESÚS JUAN

ABSTRACT

The amount of dirt or impurities in pulp and paper is a critical quality characteristic. Manufacturers evaluate dirt content by sampling a small portion of their production. In such cases, the estimate of the error associated with the measurement is as important as the measurement itself. Whether the analytical procedure is manual or automatic, the randomness inherent in the sampling process is the biggest contributor to variability. This paper presents a theoretical model that explains the observed randomness and provides an equation for estimating the error of the measurement. It is proved that if m *is the number of impurities detected in the sample, then* $1/m^{1/2}$ *is a lower bound for the relative error of the estimate and, in many cases, a good approximation of the expected error.*

THE CLEANLINESS OF A BATCH OF paper or pulp is usually evaluated by analyzing a sample of size T and measuring the area x_i of detected impurities. If m is the total number of observed impurities, the level of impurities, w, is calculated as

$$w = \sum_{i=1}^{m} x_i / T \qquad (1)$$

The measurement is usually given in units of ppm (i.e., mm^2 of dirt/m^2 of product), although it is sometimes measured as mm^2/kg of product. There are different criteria to determine the size T of the sample to be examined, depending mainly on the target pursued. (See, for example, TAPPI Test Methods T 213 om-89, "Dirt in pulp"; T 437 om-90, "Dirt in paper and paper board"; or ISO 7213, "Pulps—sampling for testing.")

Accepting the value w, obtained in a small sample, as a representative measurement of the total dirt level implies accepting a certain degree of error. Quantifying the error associated with the measurement is as important as the measurement itself.

There are various sources of error, including the inaccuracy of the measuring instruments and of the people operating them. Indeed, if the same sample is analyzed twice using the same method and instruments, different results will be obtained. The differences will be greater if the sample is reanalyzed using different methods and instruments. Nevertheless, practice has shown that these are not the main sources of error. The chief source of error is the variability among samples, a function of the random distribution of impurities in the product *(1, 2)*. Measurements obtained from different samples are different, and this is the main source of uncertainty associated with the result. Such error has the advantage of being easily measurable and monitored, and its effect is reduced when the sample size is increased.

This paper presents an equation that can be used to determine the sample size, T, required to keep the error from sampling within the desired margins. The equation is obtained from a probability model that explains the variability observed in the sampling process.

The observed differences among the dirt levels, w, of different samples are determined by two different variables:

1. The number of impurities, m, present in each sample
2. The area, x, of each impurity.

Both can be considered as random variables, and each has a very different effect on the variability (or error) that is observed in the results. The variable with the greatest potential influence—and which determines the sampling error—is m, the number of impurities detected.

The proposed model yields results equivalent to those obtained by Zeyer *et al. (2)*, as far as the equations for absolute and relative error of the estimate are concerned. However, the procedure leading to these results is different. This paper explains the influence of the two sources of error that characterize the sampling procedure: variability in the number of impurities and variability in their areas. If the observed variability in the area of the impurities is small, then the inspection system should scan a larger surface. This will increase the number of detected impurities and reduce the overall variability despite the probable tradeoff of reduced accuracy in the measurement of impurity area.

The only hypothesis of the model is that the impurities in the sample are randomly distributed according to a Poisson process *(3–6)*. This hypothesis is reasonable, taking into account the physical characteristics of the phenomenon *(3)*, and it can be easily tested applying the known procedures of goodness of fit (Chi-squared test), which are readily available in statistical software. Finally, the proposed model presents a theoreti-

NUMBER OF IMPURITIES

Sheet No.*	E_1 (0.8–1.2 mm²)	E_2 (0.3–0.8 mm²)	E_3 (0.1–0.3 mm²)	E_4 (0.04–0.1 mm²)	E_{total}	Area of impurities, mm²
			ANALYST A			
1	0	1	4	0	5	1.35
2	0	0	4	1	5	0.87
3	0	0	2	2	4	0.54
4	0	1	3	0	4	1.15
5	0	0	5	2	7	1.14
6	0	0	2	0	2	0.40
7	0	1	1	1	3	0.82
8	0	0	1	1	2	0.27
9	0	0	5	0	5	1.00
10	0	1	3	1	5	1.22
11	0	1	3	1	5	1.22
12	0	0	6	1	7	1.27
13	0	0	5	2	7	1.14
14	0	1	3	1	5	1.22
15	0	1	3	0	4	1.15
16	0	0	5	0	5	1.00
17	0	0	2	1	3	0.47
18	0	0	1	1	2	0.27
19	0	0	2	0	2	0.40
20	0	0	4	2	6	0.94
21	0	0	3	1	4	0.67
22	0	0	2	1	3	0.47
23	0	0	3	2	5	0.74
24	0	1	1	1	3	0.82
			ANALYST B			
25	0	1	3	6	10	1.57
26	0	1	4	2	7	1.49
27	0	0	2	4	6	0.68
28	0	0	1	1	2	0.27
29	0	0	1	2	3	0.34
30	0	0	4	2	6	0.94
31	0	0	1	2	3	0.34
32	0	1	3	2	6	1.29
33	0	0	2	2	4	0.54
34	0	1	4	1	6	1.42
35	0	0	4	3	7	1.01
36	0	1	2	3	6	1.16
37	0	1	3	0	4	1.15
38	0	0	5	2	7	1.14
39	0	0	1	2	3	0.34
40	0	0	1	1	2	0.27
41	0	0	1	3	4	0.41
42	0	1	3	3	7	1.36
43	0	1	4	3	8	1.56
44	0	0	3	2	5	0.74
45	0	1	1	2	4	0.89
46	0	1	1	1	3	0.82
47	0	1	1	3	5	0.96
48	0	0	2	3	5	0.61

ANALYST C

49	0	0	4	5	9	1.15
50	0	0	2	1	3	0.47
51	0	1	3	1	5	1.22
52	0	0	4	2	6	0.94
53	0	0	3	1	4	0.67
54	0	0	3	2	5	0.74
55	0	0	3	3	6	0.81
56	0	1	3	1	5	1.22
57	0	1	3	2	6	1.29
58	0	1	4	0	5	1.35
59	0	0	4	2	6	0.94
60	0	1	3	1	5	1.22
61	1	0	5	1	7	2.07
62	0	1	6	1	8	1.82
63	0	0	4	1	5	0.87
64	0	0	4	0	4	0.80
65	0	0	1	1	2	0.27
66	0	0	3	3	6	0.81
67	0	0	4	1	5	0.87
68	0	1	6	2	9	1.89
69	0	0	1	2	3	0.34
70	0	0	5	1	6	1.07
71	0	1	7	2	10	2.09
72	0	0	1	1	2	0.27

ANALYST D

73	1	1	4	0	6	2.35
74	0	1	4	1	6	1.42
75	0	0	2	2	4	0.54
76	0	0	2	0	2	0.40
77	0	0	2	2	4	0.54
78	0	0	3	2	5	0.74
79	1	2	3	0	6	2.70
80	0	2	3	2	7	1.84
81	0	1	0	1	2	0.62
82	0	1	3	2	6	1.29
83	0	2	2	3	7	1.71
84	0	1	1	2	4	0.89
85	0	0	3	0	3	0.60
86	0	0	3	2	5	0.74
87	0	0	5	4	9	1.28
88	2	0	2	2	6	2.54
89	0	0	3	4	7	0.88
90	0	0	3	1	4	0.67
91	0	2	4	1	7	1.97
92	0	0	5	1	6	1.07
93	0	1	2	1	4	1.02
94	0	1	1	3	5	0.96
95	0	1	3	0	4	1.15
96	0	0	4	1	5	0.87

* Scanned area in each sheet is 0.5 m^2

I. Impurities observed in 96 sheets distributed among four analysts

cal basis for studying and evaluating control procedures of the manufacturing process, classification techniques for product quality, bias in the estimate because of failure to detect smaller particles, etc.

PROBABILITY DISTRIBUTION MODEL

Suppose that the cleanliness analysis of a pulp or paper sample of size T reveals that M impurities have been observed with areas X_i, where $i = 1, \ldots, M$. (Capitals are used to distinguish the random variables M and X_i from the specific results observed in a sample, m and x_i.) The usual measure to indicate the batch quality is the random variable W (defined in Eq. 1), which is a function of the random quantities X_i and M. The purpose of the immediate discussion is to determine the variance of the variable W, since the square root of variance corresponds to the standard error of W.

If it is assumed that the spatial distribution of impurities in the product follows the Poisson model, then the probability that the number of impurities M appearing in a sample of size T is equal to m is

$$\Pr(M = m) = e^{-\lambda T}[(\lambda T)^m/m!] \quad (2)$$

The probability depends on a single parameter, λ, the intensity parameter. This λ represents the expected value, in this case the mean number of observed impurities per unit of surface area. If the sample has a surface area of T, then the mean impurities in a sample will be λT, i.e., the expected value of M is

$$E(M) = \lambda T$$

The Poisson distribution defined in Eq. 2 has been previously applied to the dirt-counting problem *(4–6)*. A complete explanation of this distribution and its adaptation to a specific process can be found in the literature *(3)*.

There is no need to impose any specific distribution model on the random-variable set X_i. The only required condition for the development that follows is that of independence. Both the mean, $\mu = E(X_i)$, and the variance, $\mathrm{Var}(X_i) = \sigma^2$, are unknown. The three parameters λ, μ, and σ^2 represent different characteristics of the impurity distribution:

- λ = mean number of sample impurities per sample unit
- μ = mean area of impurities
- σ^2 = dispersion of areas of impurities.

A value of $\sigma^2 = 0$ would imply that all the impurities have the same area.

The mean and variance of W in terms of the parameters λ, μ, and σ^2 are, respectively,

$$E(W) = \lambda\mu \quad (3)$$

and

$$\mathrm{Var}(W) = (\lambda T)(\mu^2 + \sigma^2) \quad (4)$$

These concepts are discussed more fully in the Appendix at the end of this paper.

Equation 4 illustrates a simple and interesting relationship. The variance, inaccuracy, or error in the estimate when using W (i.e., the variability of impurities observed in different samples of the same size T) increases with increasing λ, μ, and σ and decreases with increasing T, the sample size or scanned area. Thus as T increases, the estimate improves. W is the estimator of $W_0 = \lambda\mu$, the amount of dirt in the product. The same level of impurity can be obtained by combining different values of λ and μ. Thus many small impurities can yield the same dirt level as a few big particles. Nevertheless, these two compositions provide a different error value. It is simple to verify that for $\lambda\mu$ = constant, the variance decreases when λ increases and, therefore, when μ decreases. This implies that, from the point of view of the sampling error, the value W obtained in a process with many small particles is more accurate than that same value W for a lower number of bigger particles. The standard error (SE) of the estimate equals the square root of the

	A	B	C	D
Number of impurities in four size ranges				
E_1 (0.8–1.2 mm²)	0	0	1	4
E_2 (0.3–0.8 mm²)	8	11	8	16
E_3 (0.1–0.3 mm²)	73	57	86	67
E_4 (0.04–0.1 mm²)	22	55	37	37
E_{total}, mm²	103	123	132	124
Statistical summary of impurities				
Mean area ($\bar{x}$), mm²	0.199	0.173	0.191	0.232
Standard deviation (s), mm²	0.115	0.133	0.130	0.201
Level of impurities (w), ppm	1.712	1.775	2.099	2.399
Std. error of w [SE(w)], ppm	0.194	0.202	0.221	0.285
Relative error of w [ERR(w)]*	0.113	0.114	0.105	0.118

*ERR(W) = SE(W)/W

II. Summary of data presented in Table I for Analysts A–D

Source of variation	Sum of squares	Degrees of freedom	Mean square	*F*-ratio	Critical level
Analysts	1.817	3	0.605	2.38	0.074
Samples	23.335	92	0.253		
Total	25.152	95			

III. Analysis of variance for data in Table I

variance. Therefore,

$$SE(W) = [(\lambda T)(\mu^2 + s^2)]^{1/2} \quad (5)$$

Another way of indicating the accuracy of the estimate is by the relative error (ERR), the quotient of the standard error and the mean, also called coefficient of variation:

$$ERR(W) = \{[1 + (s/\mu)^2]/\lambda T\}^{1/2} \quad (6)$$

Equation 6 shows that as the mean impurity number decreases (λ), the relative error increases.

The previous equations are similar to those obtained in the literature *(2)*, but they were obtained following a different path. The development proposed by this paper is more detailed, since it presents a probability model to explain the distribution of impurities. As stated previously, the two sources of sample variability in the equations for estimating the error are the number of impurities, m, and their areas, x.

ESTIMATION OF ERROR

The estimation of the error is determined as a function of the parameters λ, μ, and s^2, which are unknown and must be estimated from the observed data in an inspection. If (x_1, x_2, . . . , x_m) are the areas of the m impurities observed in the inspection of a sample of size T, then

$$\hat{\lambda} = m/T$$

$$\bar{x} = \sum_{i=1}^{m} x_i / m$$

and

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

are the usual estimators for λ, μ, and s^2, respectively. An alternative for the estimator of the variance is

$$\hat{s}^2 = \sum_{i=1}^{m} (x_i - \bar{x})^2 / (m - 1)$$

When m is low, $\hat{s}^2$ offers some advantages over s^2. For high values of m ($m > 30$), the differences are insignificant. For a general description of the estimation methods, any book or statistics manual can be consulted *(7)*.

The variable $\hat{\lambda}$ is the maximum-likelihood estimator for the parameter λ of the Poisson distribution parameter. The maximum-likelihood procedure can only be applied to estimate μ and s^2 if the distribution function of the variable X_i is known. The estimators $\bar{x}$ and s^2 are the maximum-likelihood estimators when X_i has a normal distribution.

With the data from the sample, the impurity level is w, as defined in Eq. 1. The estimation of the relative error that this measurement presents is obtained by substituting the parameters for their estimators—$\hat{\lambda}$, $\bar{x}$, and $\hat{s}^2$—in Eq. 6.

$$ERR(w) = \{[1 + (\hat{s}/\bar{x})^2]/m\}^{1/2} \quad (7)$$

This equation corresponds to Eq. 9 in the work of Zeyer *et al. (2)*, with m in the divisor instead of $m + 1$.

The absolute error can also be obtained by Eq. 5. If s^2 is used instead of $\hat{s}^2$ to estimate s^2, then

$$\bar{x}^2 + s^2 = \sum_{i=1}^{m} x_i^2 / m$$

and the standard error corresponding to w can be written as

$$SE(w) = \sqrt{\frac{\sum_{i=1}^{m} x_i^2}{T}} \quad (8)$$

Dividing by another equation for the relative error is obtained:

$$ERR(w) = \frac{\sqrt{\sum_{i=1}^{m} x_i^2}}{\sum_{i=1}^{m} x_i} \quad (9)$$

When the values x_i are very similar (which is the same as saying that $\hat{s}/\bar{x} \approx 0$), the following equation is obtained from Eq. 7.

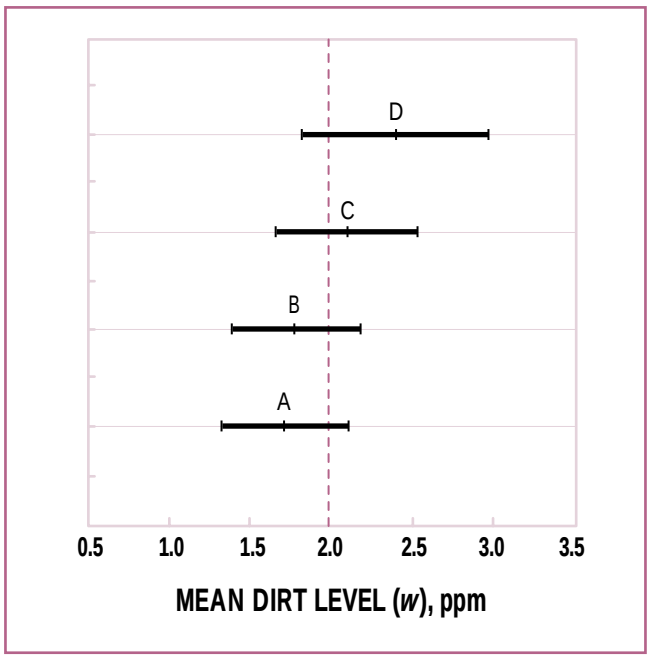
$$ERR(w) \approx 1/m^{1/2} \quad (10)$$

Since $\hat{s}/\bar{x}$ is always positive, the previous value is a lower bound, i.e., $ERR(w) \geq 1/m^{1/2}$. With a sample where 10 impurities have been detected, the relative error is higher than 31.6%, whereas with 100 particles, the bound is reduced to 10%.

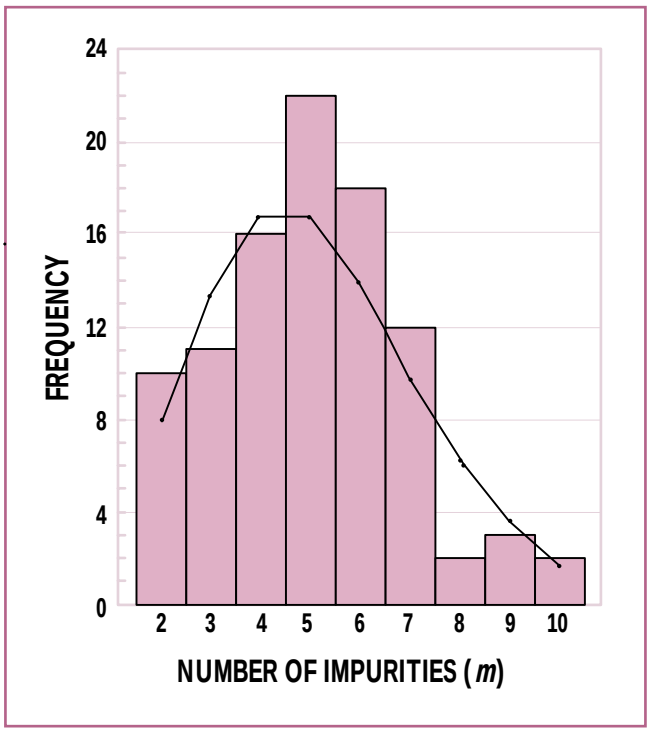
DETERMINATION OF SAMPLE SIZE T

Generally, it is important to determine the sample size T to be examined to obtain a measurement with a specific fixed error. This can be done using Eq. 6.

Let e_r be the maximum relative error expected in the estimate of the impurity level. If this is the case, then



1. Confidence intervals for mean dirt levels observed by four analysts



2. Histogram of the total number of impurities observed in each sheet. The curve depicts the expected frequencies of **m** *according to a Poisson model.*

ERR(W) $\leq e_r$, and T must comply with the following condition:

$$T \geq [1 + (s/\mu)^2]/(\lambda e_r^2) \qquad (11)$$

From Eq. 11, it can be deduced that, once the error e_r has been determined, the sample size increases when the ratio s/μ increases and decreases when λ increases. For instance, a process with twice the mean impurity number per square meter as another requires half of the sample size to have an estimate with the same relative error (assuming that s/μ is the same in both processes). In practice, λ is the predominant factor when calculating the error and determining the sample size. This implies that the number of impurities is more important than their areas in the results concerning the estimate error.

Taking into account Eq. 3 and using $W_0 = \lambda\mu$ (the level of impurities in the product), a more useful equation can be obtained by substituting $\lambda = W_0/\mu$ in Eq. 11.

$$T \geq [\mu/(W_0 e_r^2)]\,[1 + (s/\mu)^2] \qquad (12)$$

When $s = 0$, i.e., when all the particles have the same area, the sample size should verify that

$$T \geq \mu/W_0 e_r^2 \qquad (13)$$

The sample size increases proportionally with the mean area of the impurities (μ) and decreases with the amount of dirt (W_0). For nonzero values of s, Eq. 13 provides a lower bound of the sample size. For example, to estimate a mean value of $W_0 = 2$ ppm with a 10% relative error, when the mean area of the particles is 0.1 mm^2, it is necessary to take a sample bigger than 5 m^2. To reduce the error to 5%, the sample should be bigger than 20 m^2. If the level of impurities is $W_0 = 5$ ppm, then the sample sizes drop to 2 m^2 for 10% relative error and 8 m^2 for 5% error.

The previous equations imply that, given the relative error, it is not possible to recommend a unique value of the sample size. This value depends on the values of the parameters that one needs to estimate. Therefore, to fix the sample size T, it is necessary to have some *a priori* information about the parameters W_0 (or λ), μ, and s. Having an estimate of these, the sample size is established by simply substituting them in the previous equations. When there is no initial information about the value of these parameters, an initial sample of reduced size, T_0, must be taken. With this information, a first estimate of the model parame-

KEYWORDS
Dimensions, dirt count, image analysis, impurities, mathematical models, measurement, Poisson ratio, precision, pulps, random error, samples, sampling, statistical distribution, TAPPI test methods.

ters is obtained, and then an approximate value of the necessary sample size T could be obtained.

EXAMPLE

As an example, the proportion of impurities in TCF (totally chlorine free) bleached eucalyptus pulp (ENCE S.A., Spain) was analyzed. From the manufactured product, 96 sheets were chosen at random during a period of homogeneous production conditions. These sheets were distributed among four analysts, who manually evaluated the cleanliness. The central square of 0.5 m x 0.5 m was analyzed for each sheet on both sides.

Each analyst placed the sheet on an illuminated table, counted the number of impurities, and classified them according to their area. The results obtained by each of the four analysts (A, B, C, and D) are presented in Table I. Each line in the table corresponds to a surface of 0.5 m^2, and the total surface analyzed by each analyst was 12 m^2.

As can be expected, the results are different for each analyst, since there are two sources of randomness. (The number of impurities found by each analyst is different, and the area of these impurities is also different.) The results in Table II show that Analyst A found the level of impurities to be 1.712 ppm, whereas D obtained 2.399 ppm. The absolute error, SE(w), associated with each analyst was obtained using Eq. 8. This error increases when the level of impurities increases. The relative errors, ERR(w)—quotients between absolute errors and levels of impurities—are very similar, close to 11%. A lower bound can be obtained using the rule $1/m^{1/2}$, i.e., approximately 9%.

Table III shows the results after analysis of variance for the data from the four analysts in Table I. The critical level of the test is 0.074, showing that with the usual confidence levels (95%, 99%), the differences among the means for the four analysts are not significant. Only 7% (r^2 = 1.81/25.15) of the total variability is due to the differences among the analysts, whereas the remaining 93% is due to differences among the samples.

Figure 1 shows the confidence intervals (95%) for the mean level of impurities (ppm) observed by each analyst. The global mean of the 96 sheets presents a mean value of 1.99 ppm (vertical line in the figure). This value falls within the confidence intervals for the four analysts.

Finally, we tested the hypothesis that the random variable "number of impurities per surface unit" follows a Poisson distribution. The histogram in Fig. 2 shows the frequencies observed for the total number of impurities observed in each of the 96 sheets (E_{total} in Table I). The expected numbers according to the Poisson model (Eq. 2, with $\hat{\lambda}$ = 1.99 ppm and T = 0.5 m^2) appear as an overprinted curve. The goodness-of-fit test gives a critical level equal to 0.35, which implies that the Poisson model is not rejected and can be applied.

CONCLUSIONS

This paper presents a probabilistic model that explains the variability for dirt counts in pulp and paper samples. The model provides a formula for estimating the error associated with the measurements.

The variability observed in the samples is caused by two factors:

1. Each sample contains a different number of impurities.
2. The areas of the impurities in the samples are different.

Of these two sources of variability, the first is most important. The model demonstrates that the inverse of the square root of the number of impurities is a very good approximation for the relative error of the dirt estimate. Consequently, the accuracy of the estimated impurity level increases with an increase in the number of detected impurities. For two samples with an equivalent area of impurities, the estimate based on a large number of small impurities will be more accurate than another estimate based on a small number of large impurities. The only way to reduce the sampling error is by increasing the sample size to increase the number of observed particles. The model provides a method of calculating the sample size required to produce an acceptable level of error.

Accurate measurement of the area of each impurity is only relatively important. Therefore, it is better to design a measuring system that can detect a large number of impurities (i.e., by scanning a greater surface area), even at the cost of losing accuracy in the measurement of each impurity's area. For instance, better results would be obtained by increasing the scanned area and reducing the frequency of area measurement to perhaps one out of ten observations.

The model presented in this paper could also be used to design control systems, to evaluate the bias in the estimate when small impurities are not taken into account, and to design and evaluate classification procedures for the manufactured product. TJ

APPENDIX: MEAN AND VARIANCE OF W

The distribution of the variable W corresponds to a hierarchical probability model *(7)*. Such models allow the analysis of complex processes through a sequence of simple models. Given $M = m$, the number of impurities found in the sample, $W|M$, which is termed a conditional random variable, has as mean

$$E(W|M) = \mu M/T$$

and as variance

$$Var(W|M) = \sigma^2 M/T^2$$

Since M, with a Poisson distribution (Eq. 2), has as mean $E(M) = \lambda T$ and variance $Var(M) = \lambda T$, it is easy to calculate $E(W)$ *(7)*:

$$E(W) = E(E(W|M)) = \lambda\mu$$

Applying

$$Var(W) = E(Var(W|M)) + Var(E(W|M))$$

Eq. 4 is obtained:

$$Var(W) = (\lambda T)(\mu^2 + \sigma^2) \quad (4)$$

Juan is an assistant professor at the Universidad Politécnica de Madrid, C/ José Gutierrez Abascal 2, 28006 Madrid, Spain. E-mail: jjuan@ccupm.upm.es

This work was developed for the Empresa Nacional de Celulosas (ENCE S.A., Spain). The author is grateful to ENCE associates Gabriel Toval, R&D manager, and Javier Romero, research scientist, for helpful comments and for providing the data used to illustrate the concepts presented in this article. This work was supported by the Commission of the European Communities (the Specific RTD Programme), contract No. FAIR-CT95-0560 (DG 12-SSMA).

Received for review July 25, 1995.

Accepted Nov. 27, 1995.

LITERATURE CITED

1. Sutman, F. J., *Tappi J.* 77(5): 179(1994).
2. Zeyer, C., Venditti, R. A., Puangchinda, K. W., *et al.*, *Tappi J.* 78(2): 168(1995).
3. Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. I, Wiley, New York, 1968.
4. Trépanier, R., *Tappi J.* 72(12): 153(1989).
5. Jordan, B. D., *TAPPI 1992 Image Analysis Workshop Notes*, TAPPI PRESS, Atlanta.
6. Jordan, B. D., *TAPPI 1993 Process and Product Quality Conference Proceedings*, TAPPI PRESS, Atlanta, p. 25.
7. Casella, G. and Berger, R. L., *Statistical Inference*, Wadsworth, Belmont, CA, 1990.