wcsb7 proceedings

doi: 10.1255/tosf.47

The advantages and pitfalls of conventional heterogeneity tests and a suggested alternative

Francis F. Pitard

Francis Pitard Sampling Consultants, 14800 Tejon Street, Broomfield, CO 80023, USA. E-mail: fpsc@aol.com

Heterogeneity Tests have been very popular for the last 30 years and there are several versions of them such as the method of choice used by François-Bongarçon to quantify and minimize QFE_1 which is a combination of the Fundamental Sampling Error and the Grouping and Segregation Error and sometimes Analytical Error. A more recent version called "segregation free analysis for calibrating the constants K and x" is used by Minnitt, and an older, obsolete version using fragments collected one by one at random from several size fractions to calibrate the constant K was used a long time ago by Gy and Pitard. All these methods have their merits and pitfalls. The common pitfall is that they all depend on the collection of a representative composite sample consisting of about half a ton of material. In Mineral Processing it is well known how difficult it is for geologists to provide a representative sample from a given geological unit to perform reliable metallurgical testing; the same difficulties are encountered in performing Heterogeneity Tests. Furthermore, experience clearly shows that for trace constituents such as gold, many tons should be collected to obtain a reliable composite. Perhaps there is a more representative way to collect the information necessary to calculate the variance of the Fundamental Sampling Error *FSE*, which can support and complement the method of choice referred to earlier. This paper suggests that all the necessary information can be obtained by slightly modifying the logging practices of geologists. From such observations, reliable histograms of the size distribution of particles of the mineral of interest can be made representing the properties of an entire geological unit. Such information can be obtained at an early phase of exploration leading to an unmistakable definition of the sampling constant K, and possibly an accurate definition of the mathematical model of the liberation factor leading to the constant x; using modern microscopy the mineralogist can define the evolution of the liberation factor as a function of increasing comminution better than anyone else. Furthermore, this paper suggests that the determination of the liberation factor is no longer a critical factor, though most certainly useful, if using the information from modified logging practices and two old formulas suggested by Gy in the 50's instead of his famous formula using the liberation factor.

Introduction

The following material should not be perceived as a replacement for the method of choice to estimate constants K and x in a formula suggested by François-Bongarçon to quantify the variance of the short-range Quality Fluctuation Error QFE_1 affecting splitting processes in routine sampling and subsampling protocols. The suggested calibration was approved as a method of choice in a common publication by Pitard and François-Bongarçon (2011)[1] and should remain so as far as sampling practitioners are concerned. However, Heterogeneity Tests are far from perfect and no matter how careful practitioners are, there are pitfalls that can be prevented by paying attention to arguments presented in this paper. Therefore the only objective is to suggest to geologists and mineralogists that early on they can provide valuable information by adding the necessary observations on the drilling log and by making simple mineralogical tests. The added information can help to prevent great mistakes in the ways exploration and grade control data are looked at. But first some paradigms that are well accepted by sampling experts should be eliminated.

Poisson Processes and liberation issues

In the mind of many people a Poisson Process cannot take place unless the constituent of interest particles (e.g., gold particles) are liberated from their rock matrix; there is nothing that can be so far from the truth. Such belief is based on the fact that gold particles should be randomly distributed, but obviously they are not. There are plenty of geological explanations for in-situ gold particles to be distributed in a certain area of a small ore block (e.g., 15 × 15 × 15 meters). Therefore, someone may rightly object to using a Poisson model which is the simplest possible and most random way in which we may explain why the gold particles are where they are. However, all this assumes some a priori knowledge of the regionalization within that little block. We may have some of that knowledge between blocks, but not necessarily within any given block.

Therefore, before going any further, we must elaborate on the paradigm of being an observer, since the observer has no idea where the gold in that block is. He may know there is gold, but he does not know where. The resulting effect is that when he drills that block, and within that block there may be 1, 2, 3, or more clusters of gold particles somewhere, the location to drill chosen by the observer who knows nothing ahead of time is a random process of its own, even if the gold is not strictly distributed at random. So, the resulting gold content of that core, within that block, can be assimilated to a random process, not because of the way it is distributed in the deposit, but because of what the observer is doing with the selected location and selected basic volume of the support of observation as he becomes a participant; there is a subtle difference. It is exactly the same thing for coarse fragments in which the gold is not necessarily liberated.

For the purist who rightly insists that random variables be defined by reference to an appropriate probability block it is not much of a loss to take the Poisson model as a good tool to help us, especially when the observer is personally responsible for introducing the Poisson Process in the first place (i.e., no a priori knowledge and an extremely small support volume).

For anyone who may have the desire to better understand what is meant by "Poisson Process" Kingman's book, 1993, is an excellent one[2].

With this knowledge in mind it is not difficult to demonstrate that collected data for most trace constituents, and gold is one of them, are affected by a Poisson Process of some kind that was originally introduced because of the limited volume of the drilling support. This has huge implications in the collection of representative samples to perform Heterogeneity Tests and ultimately calibrate constants *K* and *x*.

Poisson Processes and trace constituent size-grade trends

To simplify the discussion let's take the example of gold as a trace constituent, keeping in mind it is applicable to any trace constituent. When there is a large in-situ nugget effect combined with a clustering effect of the gold particles that don't like to be alone, the distance between clusters within mineralization increases. Many samples do not contain coarse gold as they should. Later on, when split duplicates are taken, and ultimately fire assay duplicates are taken, since they had no coarse gold to begin with, they give the illusion that low grade samples represent areas where no coarse gold is present, even though it is true for samples coming from areas where indeed there is no coarse gold. Again, the observer will not know the difference. So, let's not feel safe by saying the low grade material does not contain coarse gold; it most certainly will in a substantial amount of cases! This deserves further thinking from geologists, grade control engineers, and geostatisticians. The author became sure of such a property in several projects where grade control vastly underestimated the gold grade going to the plant for no apparent, good reason. The same problem was clearly observed for arsenic, molybdenum and cobalt minerals. It was even observed for impurities such as silica and aluminum in iron ore, for sulfur in coal, for silica in bauxite, etc… Basically, the problem is not rare.

Indirect implications for heterogeneity tests

The appropriate approach for conducting Heterogeneity Tests for major and some minor constituents has been well established and the objective of this paper is not to question this at all. However, for low grade gold deposits for example, the conventional approach may indeed work well for deposits with finely disseminated gold, but it may be misleading when gold particles are large (e.g., superior to a few hundred microns, or when fine gold or any other trace constituent is clustering). The author witnessed many such cases and clearly there is a need to suggest a strategy to make sure sampling practitioners are not reaching over-optimistic conclusions. Again, without understanding how Poisson Processes may take place, the following material may seem bizarre for the reader.

In a letter criticizing François-Bongarçon's work Smee and Stanley (2005)[3] said "Gy's formula is based on and derivable from the binomial theorem. Consequently, Gy's formula doesn't apply to samples containing very low concentrations of elements contained in rare grains (e.g., Au, PGE, diamonds, etc.), where a Poisson relation is applicable. Our avoidance in referencing Gy stems directly from the fact that we consider samples containing nuggets to be a scenario that is inconsistent with Gy's approach." This statement shows sampling practitioners in the world of Measurement Uncertainty vastly misunderstand Gy's work and have no idea about the many subtleties of his propositions and therefore they are in no position to criticize those who apply his work in a wise and knowledgeable way.

First, the Poisson model is a limit case of the Binomial Model use by Gy, and therefore a close "cousin" and Gy was perfectly aware of nugget problems. Nobody who is knowledgeable enough would use Gy's general formula to calculate the variance of *FSE* for a sample mass that is too small by several orders of magnitude. However, anyone can turn the formula around and calculate the necessary sample mass that is required to prevent the introduction of a Poisson Process, a domain for which the formula is perfectly applicable. This is exactly what Gy always did and it is what is suggested in this paper.

Cardinal Rule #1 in sampling

Biases in sampling are the worse misfortune that may take place, and were the driving force to establish the many rules of sampling correctness, so theoretical developments of equi-probable sampling made by Gy and Matheron could apply in practice. This led to the many advances to minimize Increment Delimitation Error, Increment Extraction Error, Increment Preparation Errors and Increment Weighting Error which are the biggest contribution of Gy's theory by far according to his own words. Is this sufficient to prevent sampling biases? The answer is no. For example, it is well known that the content of a constituent of interest may drastically change from one size fraction to another. Then, plain logic would suggest the following Cardinal Rule in sampling should never be broken up: a sample mass that is too small to well represent all size fractions cannot provide a sample representative of anything else; this has huge implications for any kind of Heterogeneity Test.

Successive stages of sampling and sub-sampling may each require compliance with a pre-established limit that highly depends on the practitioner's objectives as suggested by Pitard (2013)[4]. But, the most difficult size fraction to properly represent in the sample is obviously the one containing the largest fragments. This strongly suggests some long forgotten formulas from Pierre Gy should be brought back to the rescue and a careful discussion should follow. Let's be clear, without a good understanding of these formulas there is no possible understanding of Gy's subtle work.

Gy (1971)[5] and Pitard (1993)[6] derived the following formula to calculate the variance of Fundamental Sampling Error to be used to make sure a given size fraction is well represented in collected samples.

$$s^2_{FSE} = \left[\frac{1}{M_S} - \frac{1}{M_L}\right] f \cdot \rho \left[\left(\frac{1}{a_{Lc}} - 2\right) d^3_{FLc} + \sum_x d^3_{FLx} \cdot a_{Lx}\right] \tag{1}$$

Notations are:

L_C a size fraction of interest

a_{Lc} the proportion of L_C in the lot L

M_S the mass of the collected sample

M_L the mass of the lot to be sampled

F_{Lc} the average fragment of the size fraction L_C

d_{FLc} the size of the average fragments in the size fraction of interest

d_{FLx} the size of the fragments in the other size fractions besides the one of interest

f a fragment shape factor

ρ the average density of the fragments

This formula can often be simplified for many applications:

- If $M_L > 10\ M_S$
- If d_{FLc} is not much different from d defined as the size opening of a screen that would retain 5% of the material by weight.

- If a_{Lc} is small, then

$$s^2_{FSE} = \frac{f \cdot \rho}{M_S}\left[\frac{1}{a_{Lc}} - 2\right] d^3_{FLc} \tag{2}$$

and if $d_{FLc} = d$

$$s^2_{FSE} = \frac{18 \cdot f \cdot \rho \cdot d^3}{M_S} \tag{3}$$

This convenient formula provides a filter to make sure the exponent x for d is not abused when used in a formula like one suggested by François-Bongarçon:

$$s^2_{QFE1} = \left[\frac{1}{M_S} - \frac{1}{M_L}\right] K \cdot d^x \tag{4}$$

where:

$$K = f \cdot g \cdot c \cdot (d_l)^r \text{ and } x = 2 - r$$

K and *x* are the key factors to quantify in various experiments. If $x < 3$, clearly it is not an issue when the values for *d* are below 1 cm, however it can indeed become an issue for large values of *d* such as for sampling run off mine material.

Example of application

If a run off mine material has a value of 10 cm for *d* and a 1-ton sample is required to represent the coarsest fragments with an uncertainty of 15% (1s), it would be unfortunate to recommend a much smaller mass on the basis that *x* is much smaller than 3. Obviously, the value used for *K* has a big influence on the outcome of this discussion; indeed if *K* is very high it is likely that there is no problem.

Cardinal Rule #2 in sampling

The size d_M of the grains of mineral of interest, liberated or not, must play an important role in the necessary sample mass. d_M can also be a cluster equivalent when several of those grains are very close to one another within a core sample or within a larger fragment. Gy corrected for this problem in an elegant way, not always well understood by practitioners, with his liberation factor. In other words, in his original formula with $x = 3$, both concepts *d* and d_M were preserved; be aware it is no longer the case with formula [4].

Often, especially for trace constituents, it is difficult and impractical to determine the liberation factor with sufficient accuracy, and this makes some formulas vulnerable. Enormous literature has been written on this subject, the best one by François-Bongarçon (2000, 2001)[9,10]. However, it is not a must to use the conventional, favorite approach suggested by Gy's general and well-known formula. The following suggestion is pragmatic, accurate, and falls in line with Ingamells' approach; it is summarized in the three following statements:

- Use Gy's suggested approach for liberated gold when $d_{M,}$ which is d_{Au} in formula [5], becomes the dominant factor; it can be generalized to many other components of interest.

$$s^2_{FSE} = \left[\frac{1}{M_S} - \frac{1}{M_L}\right] \frac{f_{Au} \cdot g_{Au} \cdot \rho_{Au}}{a_L} d^3_{Au} \tag{5}$$

- Verify that the sample mass suggested by the generalized version of equation [4] is compatible with the mass necessary to represent all size fractions in the lot by using equation [1], or [3].
- The largest required sample mass for a pre-selected precision, obtained by equation [1] or [3] (i.e., using *d*) and equation [5] (i.e., using d_M defined below) necessarily takes priority on deciding what the sampling protocol should be.

Generalization of equation [5] by defining new notations:

f_M the shape factor of the constituent of interest

g_M the particle size distribution factor of the constituent of interest

ρ_M the density of the constituent of interest

d_M the maximum size of the constituent of interest particle, liberated or not, or cluster of such particles contained in a single fragment of the surrounding matrix; d_M is defined as the size of a screen that would retain no more than 5% by weight of all the particles of the constituent of interest.

Thus, we obtain the very useful simplified formula:

$$IH_L = f_M \cdot g_M \cdot d^3_{IM} \cdot \frac{\rho_M}{a_L} \tag{6}$$

Useful sampling nomographs can be calculated with the following formula:

$$s^2_{FSE} = \left[\frac{1}{M_S} - \frac{1}{M_L}\right] \frac{f_M \cdot g_M \cdot \rho_M \cdot d^3_M}{a_L} \tag{7}$$

The great advantage of this approach is its accuracy and the easiness to collect the relevant and necessary information through microscopic observations, and it should somewhat reconcile Gy, Ingamells, and François-Bongarçon. In the event reconciliation is not possible it should be a clear indication some heterogeneity properties of the constituent of interest are still unknown and further investigation is needed. This debate naturally leads to Cardinal Rule #3.

Another advantage of equation [7] is for subsampling finely ground material, as some constituents such as soft ones like gold, molybdenite, galena and many more do not comminute well. Very hard minerals like chromite may show the same problem. For example a sample pulverized to 99% minus 106 microns may still contain a 300-micron gold particle making all other formulas weak and perhaps misleading.

Cardinal Rule #3 in sampling

As Pierre Gy said many times, especially when criticizing the work of Richard (1908)[15], when deciding what the exponent of *d* should be, and therefore the constant *x*, there is a confusion between *FSE*, QFE_1, and even the Analytical Error *AE* poorly defined by non-chemists and TOS experts. This confusion has been responsible for over a century for total chaos, and still remains an issue today. Problems are:

1) For very fine material the variance of *FSE* rapidly becomes a negligible factor unless unrecognized delayed comminution takes place for the constituent of interest.
2) The segregation error can be huge as the constituent of interest is liberated and possibly of a very different density than the rest of the material.
3) Taking the optimistic assumption that analytical increments are taken perfectly at random (an absolute requisite for Gy's definition

of *GSE*), which is rarely the case at the balance room of a laboratory, the variance of *GSE* can become small indeed; however it takes work an analytical chemist is not willing to spend the time on. As a result, the segregation error which is no longer *GSE*, may become vastly underestimated because it no longer obeys rules set by the TOS.

4) The Analytical Error *AE* cannot be estimated by doing replicate assays that include the last *FSE* and last *GSE*. Let's assume the chemist takes a 30-g analytical subsample for fire assay; the taking of that sample has nothing to do with the Analytical Error which includes fusion, cupellation, acid digestion of a bead, contamination, losses, spectrometer calibration or use of a precision balance, additive and proportional interferences, etc… In other word it is very hard, if not impossible in some cases to appreciate what *AE* really is. Furthermore, *AE* is extremely operator dependent. There is no such things as a bad analytical method, there are only incompetent analysts who apply it for the wrong conditions.
5) There is no such thing as a segregation free analysis when taking replicate samples in a given size fraction as particles segregate even if they are all the same size. They will most certainly segregate because of density, shape, electrostatic property differences, etc…

All this is clearly summarized in the sketch illustrated in Figure 1 and very familiar to Visman, Ingamells and Gy through verbal conversations, and many others who were wise enough to admit that what they measured with replicate samples or replicate assays may have nothing to do with the variance of *FSE*. It can be noticed as well that in this figure when segregation is mentioned it is not necessarily referring to *GSE* as defined in the TOS; the subtle difference depends on what the operator may do. The only thing the author asks is not much to comply with: call variance sources by their respective name instead of calibrating *x* to compensate for things that are not clearly defined or understood. An example is appropriate: an operator shakes a laboratory pulp to collect a tiny analytical sample, then makes the assumption there is no longer any significant segregation in the pulp, and finally takes one or two tiny increments with no respect to the TOS. The resulting variance, after guessing what the analytical variance should be and removing it is found to be large. The operator put the blame on a large variance of *FSE* when it is clear that he was introducing a massive segregation variance because of the way he collected the increments. In this particular case he was introducing a variance that has nothing to do with FSE, nor GSE, because all the subtle principles clearly defined in TOS were completely ignored therefore prohibit the segregation variance to be a random one as it should be.

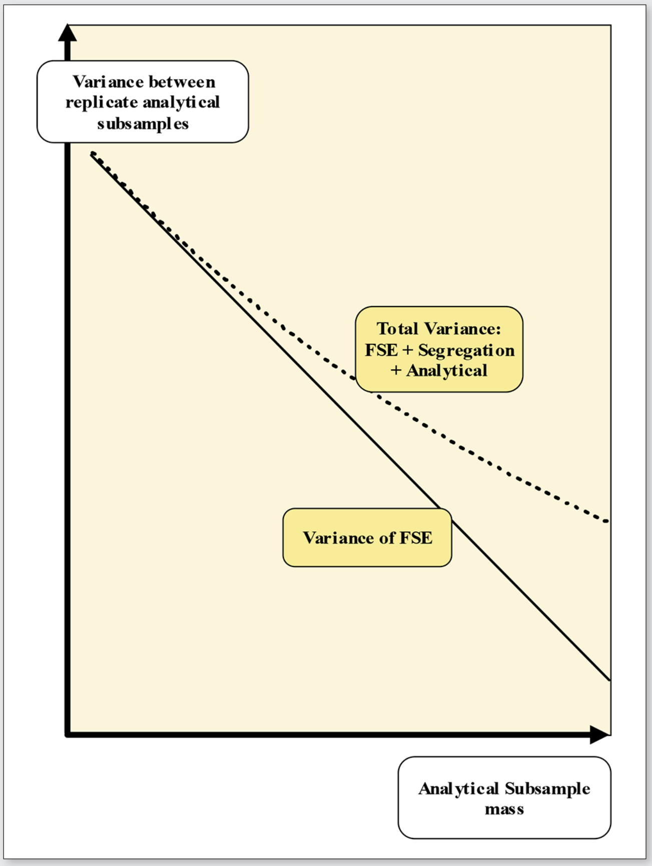

Figure 1. Replicate assays variance and its components.

Suggesting a new integrated iterative approach

Iteration is the word of wisdom in sampling. The following three steps are not necessarily suggested in chronological order. Rather, each step can be taken simultaneously which ultimately will provide confidence that no stone has been left unturned.

Step #1: The mandatory calibration of *K* and *x*

The calibration of constants *K* and *x* in equation [4] as suggested by François-Bongarçon is a mandatory step that is non-negotiable; please notice notations in that formula very carefully. Indeed, the use of the notation QFE_1 is valid only if the operator has been collecting many increments in full compliance with sampling correctness, which is a very optimistic assumption as experience proves. If not in full compliance, then the resulting variance is anyone's guess because there is no longer any theoretical development possible as demonstrated by Gy and Matheron. Such calibration allows minimizing the variance of the Fundamental Sampling Error and also measures the leftover effect of the Grouping and Segregation Error depending on the equipment used to split samples at the sample preparation room and at the laboratory, and on the operator's training which can be a huge factor. For the details of such procedure the reader is referred to François-Bongarçon's publications (2000 and 2001)[9,10].

Step #2: The geologist to the rescue

It is necessary to better log the properties of gold in each geological unit. With minor modifications the same list may easily apply to other constituents of interest in iron ore, in coal, in porphyry copper deposits, and others, gold being only a convenient example. For each core sample within substantial mineralization the following information should be carefully logged:

- Where is the gold?
- What are the associations of gold?
- How much gold is finely disseminated within sulfides, such as pyrite or other minerals?
- How much gold is coarse and perhaps nearby other minerals?
- Are gold and pyrite or other mineral occurrences associated with narrow or large quartz veins? If so, are there several quartz events?
- Study size distribution of gold particles. A good histogram is needed for each geological unit. After observing several thousand

samples within mineralization it should be possible to roughly estimate the size d_M above which only 5% of the gold can report.

- Equally important, study the size distribution of gold particle clusters; in other words when you see one gold particle (measure it), how many more gold particles are in the immediate vicinities? e.g., 10 or more within 100 cm^3? After observing several thousand samples within mineralization it should be possible to roughly estimate the size $d\text{¢}_M$ above which only 5% of the gold can report as cluster equivalents.
- Etc…

Step #3: The mineralogist to the rescue

Suarez and Carrasco (2011)[13] demonstrated in an unambiguous way that careful mineralogical studies can provide valuable information to model the variability of the liberation factor as a function of comminution stage. It is very unfortunate such study does not generate more interest. The same study suggests that the maximum content model suggested many years ago by Gy is a very reliable model that was used all the time in a mineral processing research laboratory (Minemet); see Gy (1956)[11] and Pitard (1993[6] and 2009[14]).

Step #4: The selected sample or subsample mass must fairly represent the coarsest fragments

This task is easily done by using formulas [1] or [3].

Step #5: The selected sample or subsample mass must fairly represent the largest particles of a given constituent of interest

This task is done by using formula [7]. This is critically important for constituents showing delayed comminution. Usually, soft minerals such as gold, galena, molybdenite and very hard minerals such as chromite can show such problem. As a good example, the coarse gold case shown by Pitard and Lyman (2013)[12] clearly shows that a Heterogeneity Test performed by using conventional 30-g fire assays would most likely have led to very misleading conclusions; the test is not the problem, but the completely inappropriate 30-g subsample is the issue, in other words the operators would have used the wrong tools.

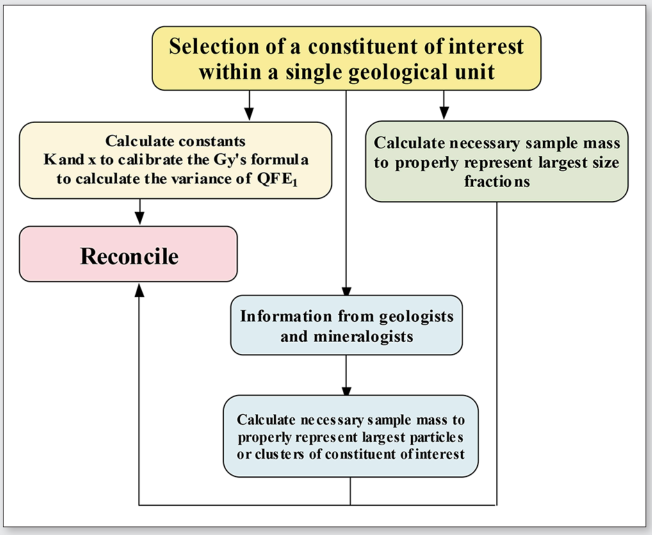

Figure 2. Logical Heterogeneity Test Flow Sheet.

Step #6: A logical flow sheet to perform Heterogeneity Tests

Figure 2 summarizes the necessary steps to perform a reliable Heterogeneity Test for various constituents of interest during exploration and grade control; the approach can easily be extended to other materials in other industries. The reconciliation box has a very important mission in cases where conclusions are grossly different: a logical explanation must be found that may lead to important decisions concerning the selection of fully optimized sampling and subsampling protocols.

References

1. Pitard F., and François-Bongarçon, D. Demystifying the Fundamental Sampling Error and the Grouping and Segregation Errors for Practitioners. Fifth World Conference on Sampling and Blending. Proceedings printed by Gecamin. 25-28 October 2011, Santiago, Chile, pp 39-55.
2. Kingman J.F.C. Poisson Processes. Oxford Studies in Probability.3. University of Bristol. Clarenton Press. Oxford, 1993.
3. Smee B.W. and Stanley C.R. Reply to Dr. Dominique François-Bongarçon, - Publication Type in a journal article: Explore - Association of Exploration Geochemists Newsletter, Volume 127, p.19-23 (2005)
4. Pitard F. Guidelines for acceptable allotted sampling uncertainty. Sixth World Conference on Sampling and Blending. Proceedings printed by Gecamin. November 19-2, 2013, Lima, Peru, chapter 2, paper 41.
5. Gy, P.M., L'Echantillonage des Minerais en Vrac (Sampling of particulate materials). Volume 2. Revue de l'Industrie Minerale, St. Etienne, France. Numero Special (Special issue, September 15, 1971, chapter 19).
6. Pitard, F.F., Pierre Gy's Sampling Theory and Sampling Practice. *Textbook published by CRC Press*, Inc., 2000 Corporate Blvd., N.W. Boca Raton, Florida 33431. Second edition, July 1993.
7. Ingamells, C.O, and Pitard F., Applied Geochemical Analysis. Wiley Interscience Division, John Wiley and Sons, Inc., New York, 1986. 733 pages textbook
8. François-Bongarçon, D. Fishy samples: how big a sample to avoid the infamous "Poisson effect"? Fourth World Conference on Sampling and Blending. The Southern African Institute of Mining and Metallurgy, 2009.
9. François-Bongarçon, D. (2000). The most common error in applying Gy's formula in the Theory of Mineral Sampling, and the history of the liberation factor. *monograph "Toward 2000", AusIMM*.
10. François-Bongarçon, D and Gy, P (2001). The most common error in applying 'Gy's Formula' in the theory of mineral sampling and the history of the Liberation factor, *in Mineral Resource and Ore Reserve Estimation – The AusIMM Guide to Good Practice*, pp 67-72 (The Australasian Institute of Mining and Metallurgy: Melbourne).
11. Gy, P.M., Poids a Donner a un Echantillon – Abaques d'Echantillonage. French edition of a former paper in German. Revue de l'Industrie Minerale, 1956, pp. 1-30.
12. Pitard F.F. and Lyman, G.J., Single and multi-stage Poisson processes: A case study for gold exploration. 6th World Conference on Sampling and Blending. Proceedings printed by Gecamin. 19-22 November 2013, Lima, Peru.
13. Suarez, E.V., and Carrasco P., Numerical studies of texture and liberation in microscopic images, Fifth World Conference on Sampling and Blending", 25-28 October 2011, Santiago, Chile. Proceedings printed by Gecamin.

14. Pitard, F.F. “Pierre Gy’s Theory of Sampling and C.O. Ingamells’ Poisson Process Approach, pathways to representative sampling and appropriate industrial standards”, doctoral thesis in technologies, Aalborg University, campus Esbjerg, Niels Bohrs Vej 8, DK-67 Esbjerg, Denmark, 2009.
15. Richard, R.H., Ore dressing, sampling. Vol.2: pp. 813-884, Vol. 3: pp. 1571 1578, Vol. 4: pp. 2031-2033. Mc Graw-Hill. New York (1908).