Click to see slides from oral presentation

[EA] Uncertainty Distributions in Practice: Which, How, What?*

Jan de Jager[1]

Search and Discovery Article #42528 (2020)**
Posted May 18, 2020

*Adapted from extended abstract based on oral presentation given at 2019 AAPG Asia Pacific Technical Symposium, The Art of Hydrocarbon Prediction: Managing Uncertainties, August 7-8, 2019, Bogor, (Greater Jakarta), Indonesia

[1]PetroEDGE, Houston, Texas (jan.dejager@hotmail.com)

Abstract

From the time the assessment of risks and volumes of undrilled prospects is done with probabilistic computer applications we have been required to specify the uncertainty of geological input parameters such as reservoir thickness, net/gross, hydrocarbon saturation, column length, etcetera. We normally do this by using some type of mathematical distribution type. Modern probabilistic tools offer a wide – and for many explorers bewildering – range of distribution types to choose from: normal, lognormal, beta, stretched beta, uniform, triangular, double-triangular, gamma, etc. In addition to these the 'good old' histogram can be used also in most tools (Figure 1).

Not all explorers have a good understanding of the available distribution types and of their differences. The workflow of the probabilistic tools that we use requires us to first choose a distribution type, and to then specify a Low value (either Minimum or P90), a High value (either Maximum or P10) and a Middle value. These inputs determine the shape of the distribution. Somewhat confusing is that the lognormal distribution requires as 'middle' input a value for the Mean, while for other distribution types it must be the Mode. It all seems to be very much about statistics instead of being about geology.

There is a common believe with many explorers that uncertainty distributions for subsurface parameters must in general be lognormal (Figure 2). This comes from observations by P.R. Rose (2001) and others (Megill, 1984; Capen et al., 1984, 1992) that "most important geotechnical parameters involved with oil and gas occurrences are lognormal". And indeed, for many geological parameters it is valid that when sufficient measured data points are plotted, they tend to display a lognormal distribution; these are so-called variability distributions: distributions of measured data. This is then taken as indication that our pre-drill uncertainty distributions (distributions reflecting the uncertainty of a precise value) for the same subsurface parameters must also be lognormal in shape. But is this correct? In Begg, Bratvold and Welsh (2014) the fundamental difference between variability and uncertainty distributions is discussed.

Uncertainty Distributions

Uncertainty distributions are representations of our degree of uncertainty of a precise single value in an undrilled prospect. Uncertainty distributions should be based on all available data, geological models and our understanding, but they are not direct representations of data. The width and shape of uncertainty distributions depend on the degree of our understanding and knowledge. As we acquire data and gain understanding, uncertainty distributions normally become narrower. This then means that there is no correct or wrong distribution; nor is there a correct or wrong distribution type. If we would have complete knowledge there would be no distribution; the parameter then becomes a single value - a spike; for example the prospect's exact average porosity, its average reservoir thickness, the precise HC column length, etc. In the subsurface there is no uncertainty; the uncertainty is in our heads. It is interesting to realize that contrary to uncertainty distributions, variability distributions may actually get wider with more data; this will be the case when the values of newly collected data are higher or lower then the extremes of the previously collected data.

It seems a basic truism that uncertainty distributions should cover all outcomes that are geologically possible and consistent with an oil or gas discovery. Nevertheless, look-back studies show that very often the actual geological parameters measured in exploration wells fall outside the predicted pre-drill uncertainty range. Clearly, uncertainty tends to be underestimated.

For uncertainty distributions it is important that they are consistent with our geological understanding of the specific geological parameter. This is not about statistical concepts, but about pragmatically translating geological understanding into realistic numbers and ranges of numbers, whilst avoiding biases.

Some simple guidelines will be presented that will allow explorers to focus on geology instead of on (possibly confusing and certainly distracting) statistical concepts. The very versatile Beta distribution, which is available in most probabilistic tools, is very convenient in this respect (Figure 3). Based on inputs of minimum, mode and maximum (or P90, mode and P10) this distribution can be skewed positively or negatively and it can be symmetrical. Therefore, using this distribution will not lead to unnecessary discussion of statistical issues; it allows geoscientists to focus on the real issue: geology and the uncertainties regarding the input parameters. The main question should always be; what is the most likely value of the input parameter (the mode), and what could the maximum be in the most optimistic case, and how low could the minimum be in the most pessimistic case.

In some cases none of the available distribution types will portray the uncertainty of some input parameters adequately. This may be when there is a high likelihood that an input value is within a specific (maybe limited) range, with a slim chance that it is much higher (or lower) than that range. With mathematical distributions and input of minimum and maximum values, the upside (or downside) will tend to be given a much too high probability. In those cases a hand-made histogram distribution can be more consistent with our understanding of the uncertainty. Several other issues that have to do with uncertainty distributions and that may lead to - sometimes passionate - discussions will be addressed also.

References Cited

Begg, S.H., R.B. Bratvold, and M.B. Welsh, 2014, Uncertainty vs Variability: What's the difference and why is it important, SPE 169850.

Capen, E.C., R.E. Megill, and P.R. Rose, 1984, Evaluating and Managing Petroleum Risk: AAPG, Tulsa, Oklahoma.

Capen, E.C., 1992, Dealing with Exploration Uncertainties, *in* R. Steinmetz, ed., The Business of Petroleum Exploration, AAPG, Tulsa, Oklahoma, p. 29-61.

Megill, R.E., 1984: An Introduction to Risk Analysis (second edition): Penwell Publishing Company, Tulsa, Oklahoma.

Rose, P.R., 2001, Risk Analysis and management of Petroleum Exploration Ventures: AAPG Methods in Exploration Series, No 12.

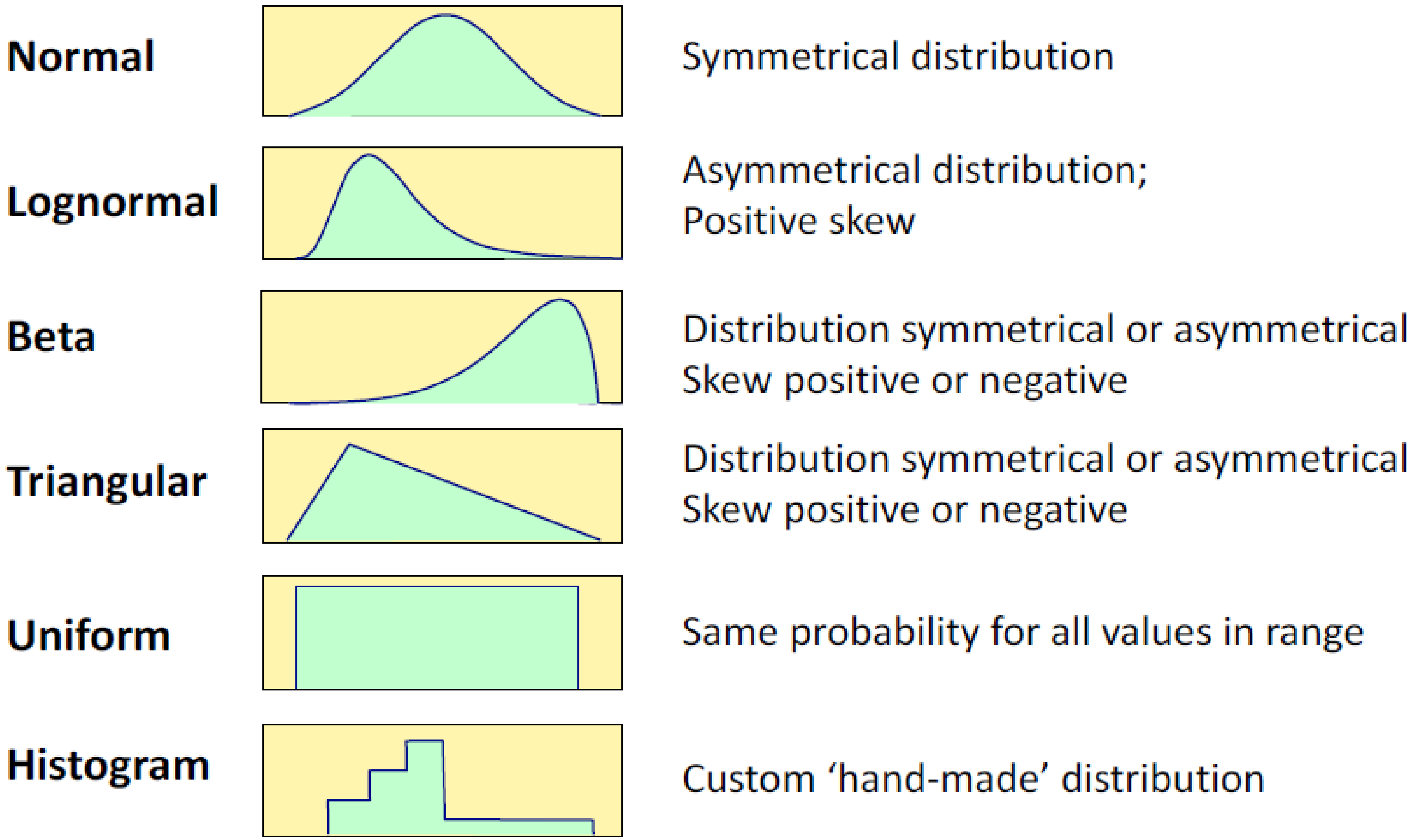

Figure 1. Modern probabilistic risk and volume tools offer a range of distribution types to indicate the uncertainty of what the average values are for the inputs for probabilistic volumes calculations, such as gross reservoir thickness, net/gross, porosity, etc.

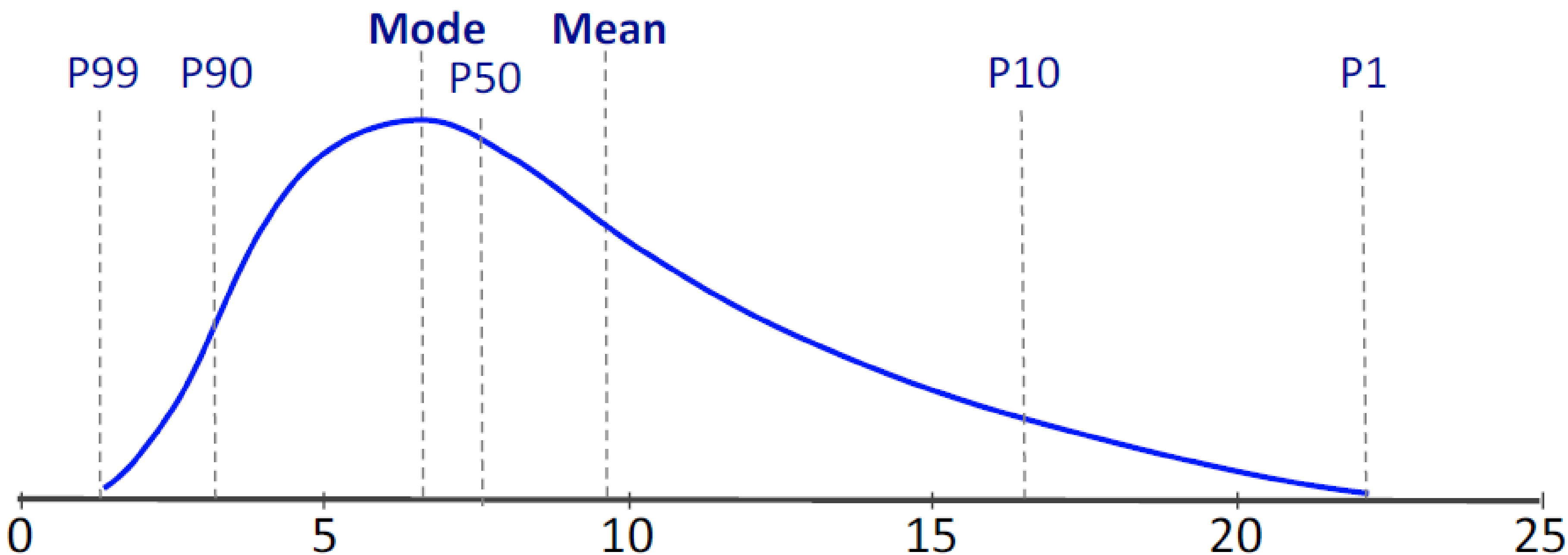

Figure 2. The lognormal distribution is often used to input the uncertainty of the average values that are used to calculate prospect oil and gas volumes. Lognormal distributions always have a positive skew, with the Mode (=most likely) much closer to the minimum value than to the maximum value. The Mean is towards the tail of the distribution, and is higher than the 'most likely' value.

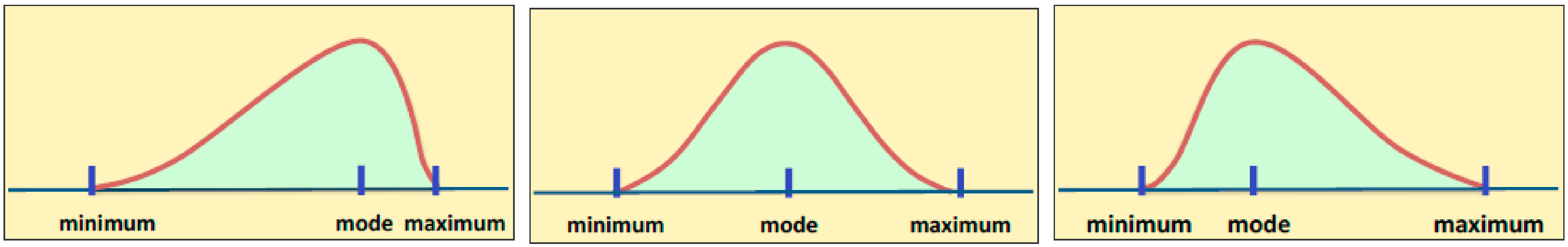

Figure 3. The versatile Beta distribution can have a negative skew, positive skew and can be symmetrical. It is a hassle-free distribution that allows focusing on the most likely value of the input parameters (the mode), and on their minimum and maximum values.