Marko Kannisto

Public
25.8.2026

CALCULATION OF SIMULTANEOUS DEFROST CYCLES FOR HVAC UNITS USING THE BINOMIAL DISTRIBUTION

Summary

The heating power used during the defrost cycles of ventilation units is typically several times greater than that used during normal operation; consequently, the building's electrical supply is often dimensioned on the assumption that all units are defrosting simultaneously. This guide describes a statistical method — the binomial distribution — used to calculate the maximum number of machines that can operate simultaneously at peak power at a given confidence level, and how this is used to derive the design power corresponding to the actual usage of the electricity supply connection. The method is generally suitable for assessing the simultaneity of several similar, intermittently operating electrical loads — such as the defrost cycles of ventilation units or the compressor cycles of heat pumps. These guidelines are intended for electrical and HVAC designers who are sizing a building's electrical connection or main fuses for a group of similar, intermittently operating loads.

1. Framing the problem

When designing the electrical systems for ventilation units, a question often arises: just how much of a load do the units' heating elements actually place on the building's electrical system? Some of the ventilation units available on the market use a lower heating element than others, which means that the unit's stated maximum output appears to be low. This may also seem attractive from an electrical design perspective, as a lower maximum power rating appears to result in a smaller electrical connection and smaller fuses.

However, the maximum power rating of a device does not tell the whole story: the maximum power rating of an individual machine and the building's actual, simultaneous electrical load are not necessarily the same thing.

Why is heating capacity required in a ventilation unit?

The purpose of the heating element is to raise the air temperature. This is a purely physical phenomenon and cannot be altered by software or control logic. The change in air temperature (ΔT) can be calculated using the following formula:

$$\Delta T = P / (q \times \rho \times cp)$$

where:

- P = heating capacity (kW)
- q = air flow (m^3/s)
- ρ = density of air (1.2 kg/m^3)
- cp = specific heat capacity of air (1.01 kJ/kgK)

For example, if the air flow rate is 50 l/s and the heating element has a power of 900 W, the air temperature will rise by approximately 15 °C. If the outside air temperature is −20 °C and the heat recovery cell is bypassed during defrosting, the supply air temperature will be approximately −20 °C + 15 °C = −5 °C.

In most situations, air this cold cannot be channelled into the flat. In practice, the options are to increase the heating output, reduce the air flow, or a combination of the two. For this reason, some manufacturers use a higher heating output than others.

The paradox of fuse sizing

However, the higher heating output raises a new question. If, for example, a block of flats has 58 flats and the maximum defrosting power of each ventilation unit is 2.0 kW, the following calculation can easily be made:

58 × 2.0 kW = 116.0 kW

On this basis, it would appear that the building's electrical system must be designed to handle a load of over 100 kilowatts for the ventilation units. But does this sort of situation ever actually happen in practice? In theory, yes. In practice, this is highly unlikely — and this guide explains just how unlikely it is and on what grounds.

2. Method / principle

Purpose and scope

This guide describes a statistical method for estimating the simultaneous peak power of several independent, intermittently operating electrical loads. The method is suitable for situations where there are several loads (n in number) that are technically identical; each load operates at peak power only intermittently during periods of known duration; and the timing of these periods is determined by load-specific, mutually independent factors. Typical applications include defrost cycles in ventilation units, compressor cycles in hot-water storage tanks and heat pumps, thermostat-controlled heating elements and other cyclically operating appliances in domestic settings.

The result of the method is the design value k: the maximum number of loads that can be operated simultaneously at peak power at the selected level of confidence. From this, the simultaneity factor k/n and the rated power of the group are derived.

Basic concepts

The utilisation rate p is the probability that, at a random point in time, a single load is in the peak power period. When the duration of a cycle is t and the interval between cycles (from the start of one cycle to the start of the next) is T, the utilisation rate is p = t / T. For example, a 20-minute defrost cycle every three hours gives p = 20/180 ≈ 11.1 per cent.

The simultaneity factor (equalisation factor, crossover factor) is a ratio used in design calculations that indicates what proportion of the combined maximum power of the loads may occur simultaneously. In electrical design, a similar concept has become established for calculating peak power (e.g. the derating factor procedures for ST circuit boards); the method set out in this guide derives the factor from the equipment's operating parameters, rather than selecting it from a table or based on experience.

Confidence level is the probability that the sum of simultaneous loads will be less than or equal to the design value k. For example, a 99.9 per cent confidence level means that, at any random point in time, the probability of the limit being exceeded is 0.1 per cent.

Marko Kannisto

Binomial distribution

When n loads operate independently of one another and each is operating at peak power with probability p at the time of observation, the number X of loads operating at peak power simultaneously follows a binomial distribution: $X \sim Bin(n, p)$.

The key figures for the distribution are:

- Expected value: $E(X) = n \cdot p$
- Variance: $Var(X) = n \cdot p \cdot (1 - p)$
- Standard deviation: $\sigma = \sqrt{(n \cdot p \cdot (1 - p))}$

The probability of a single value is $P(X = x) = C(n, x) \cdot p^x \cdot (1 - p)^{n-x}$, where $C(n, x)$ is the binomial coefficient. In design, the quantity of interest is the cumulative distribution function $P(X \leq k)$, i.e. the probability that there are at most k simultaneous loads.

Calculation of the design value using the normal approximation

When n is reasonably large, the binomial distribution can be approximated by a normal distribution. The design value at the desired level of confidence is obtained using the formula:

$k = \lceil E(X) + z \cdot \sigma \rceil$

where z is the one-sided quantile coefficient of the normal distribution and $\lceil\ \rceil$ denotes rounding up. The most common odds:

Level of certainty	**z**
95%	1.645
99%	2.326
99.9%	3.090
99.99%	3.719

The normal approximation is sufficiently accurate when $n \cdot p \cdot (1 - p)$ is at least of the order of 5–10. In cases falling below this threshold (small n or very small p), it is advisable to calculate the design value directly from the cumulative distribution function of the binomial distribution or using the Poisson approximation. The accuracy of the approximation is improved by a continuity correction: when estimating the exceedance probability $P(X \geq k + 1)$, the value $k + 0.5$ is used as the argument of the normal distribution.

At the boundaries of the validity range, the approximation typically slightly underestimates the right-hand tail of the distribution, which may result in the design value being one unit too low. The principle is important: the normal approximation is always an initial estimate, which is verified by a precise binomial calculation, and the confidence level to be reported is derived from the precise calculation — not from the approximation.

Choosing a confidence level — point-in-time vs. periodic excess risk

The confidence level applies to a single, random point in time. However, a long load cycle comprises many observation points, so the probability of an exceedance occurring during a given cycle is greater than the probability of an exceedance at any single point in time. A rough estimate: if a period contains M time windows that are approximately independent of one another (the length of each window ≈ the duration t of the peak power period) and the instantaneous probability of exceedance is α, then an exceedance will occur at least once during the period with a probability of:

P(exceedance during the period) ≈ 1 − (1 − α)^M ≈ M · α, where M · α is small.

Example: a continuous three-day period of sub-zero temperatures comprises around 200 20-minute intervals. If the instantaneous confidence level is 95 per cent (α = 5 per cent), it is virtually certain that the limit will be exceeded during the period ($1 - 0.95^{200} \approx 100$ per cent). If the confidence level is 99.9 per cent (α = 0.1 per cent), the probability of an exceedance during the period is approximately 200 × 0.001 ≈ 18 per cent — and an exceedance typically means just a single load exceeding the design value, not a collapse of the distribution.

A practical rule of thumb: select the instantaneous confidence level such that α ≈ (acceptable periodic risk of overcurrent) / M. In fuse sizing, a 99.9 per cent confidence level is generally justified, as the magnitude of the overcurrent is small (a single load), the time-current characteristics of the fuses can tolerate short-term, minor overcurrents, and the basic assumptions of the calculation (all loads at maximum density continuously) are in themselves conservative, which provides an additional margin.

The presumption of independence — the cornerstone of the method

The binomial model assumes that peak power periods for the loads start independently of one another. This is the most important and sensitive assumption underlying the method, and it must be assessed separately for each application.

A shared environment does not undermine independence. Often, all loads are exposed to the same external factor (e.g. the outside temperature). However, this does not synchronise the cycles, provided that the calculation is carried out on a 'worst-case' basis: a common factor determines the average frequency of the cycles (the value of p, which is set to reflect the worst-case scenario), but the timing of individual cycles is determined by load-specific factors. In the case of ventilation units, the timing is determined by the amount of condensation that accumulates on the heat exchanger, which depends on the humidity load in the home — the number of occupants, bathing, cooking and the times at which these activities take place. The same principle underlies the classical peak capacity calculation for telephone and electricity networks (Erlang traffic theory): the shared environment determines the level of load, whilst the timing of individual events is randomly distributed.

Independence is compromised if the trigger mechanism is shared. Typical risk situations:

- **Threshold-based activation:** if a cycle starts directly from the threshold value of a common variable (e.g. "defrosting begins when the outdoor temperature falls below −10 °C"), all loads are activated almost simultaneously when the threshold is exceeded. Period intervals of varying lengths for each load spread the phases over a few cycles, but the first synchronised start-up may still cause a design peak.
- **Time control:** time-based cycles (e.g. "defrosting every hour on the hour") are perfectly synchronised and never get out of sync.

- **Recovery following a power cut:** After the power cut, all devices start up simultaneously. If the device switches directly to peak power mode upon start-up, the instantaneous simultaneity may be 100 per cent.
- **Centralised control:** commands sent simultaneously to all devices by a building automation system or a cloud service.

Before applying the method, it must be verified via the device's control logic that the trigger is demand-based (based on load-specific measurement). If the trigger mechanism is unknown, independence cannot be assumed.

Varying utilisation rates

In reality, load utilisation rates vary (the moisture load in a studio flat ≠ that of a family of five). Strictly speaking, the number of simultaneous loads follows a Poisson-binomial distribution in this case, where each load has its own p_i: the expected value is Σp_i and the variance is $\Sigma p_i(1 - p_i)$. When the calculation uses a worst-case p value common to all loads ($p \geq$ all p_i), both the expected value and the variance are greater than the actual values — the result is therefore a genuine upper bound, not an average estimate. This simplification is both easier to calculate and safe in terms of design.

Calculation formula

- 1. Specify the parameters of the cycle: the duration t of the peak power cycle and the cycle interval T under the worst-case realistic operating conditions. Document the basis for these values (control logic, field experience, measurements).
- 2. Calculate the utilisation rate: $p = t / T$.
- 3. Ensure independence: check the release mechanism. If the condition of independence is not met, the method cannot be used as it stands.
- 4. Calculate the key figures: $E(X) = n \cdot p$ and $\sigma = \sqrt{(n \cdot p \cdot (1-p))}$.
- 5. Select the confidence level based on the exposure time; for fuse sizing, this is typically 99.9 per cent.
- 6. Calculate the design value: $k = \lceil E(X) + z \cdot \sigma \rceil$. If $n \cdot p \cdot (1-p) < 5$–10, calculate k precisely using the cumulative distribution function of the binomial distribution.
- 7. Calculate the rated power: P_mit = k × P_peak + (n − k) × P_norm, where P_peak is the load power during the peak power period and P_norm is the load power at other times. The simultaneity factor for peak power is k/n.
- 8. Document the boundary conditions: the basis for independence, exceptional behaviour (e.g. a power cut), and the status of the initial values (guaranteed limit vs. empirical estimate).

Practical calculation tools

Excel / LibreOffice:

- Design value directly: =BINOMIAL.DISTRIBUTION.INV(n; p; confidence level) (English: BINOM.INV). Example: =BINOM.DIST(58; 0.1111; 0.999) → 15.
- Checking coverage for the selected k: =BINOMIAL.DISTRIBUTION(k; n; p; TRUE) (English: BINOM.DIST) returns P(X ≤ k). Example: =BINOMIAL.DISTRIBUTION(14; 58; 0.1111; TRUE) → 99.87 per cent.

A precise manual calculation using a recursive formula. The binomial distribution can be calculated accurately using a basic calculator without binomial coefficients, as the successive probabilities are obtained by simple multiplication:

- Initial value: $P(X = 0) = (1 - p)^n$
- Recursion: $P(X = x + 1) = P(X = x) \times [(n - x) / (x + 1)] \times [p / (1 - p)]$

Calculate the values in the order P(0), P(1), P(2), ... and sum them up until the total exceeds the desired confidence level — the final x is the design value k. The workload typically takes a few minutes (15–20 multiplications). The initial value $(1 - p)^n$ can be calculated using the calculator's power key or via the logarithm, e.g. $(8/9)^{58} = e^{(58 \times \ln(8/9))} = e^{(-6.831)} \approx 0.00108$.

A normal approximation as a rough estimate using pen and paper provides a quick result without the need for spreadsheet calculations, but the final design value and the confidence level to be reported are always derived from a precise calculation (using Excel or recursion).

3. Application example

Initial data: A block of flats with 58 flats.

- n = 58 ventilation units.
- Duration of the defrost cycle: t = 20 mins.
- Defrosting time under the worst-case known operating conditions: T = 180 min.
- The machine's maximum power during defrosting is P_peak = 2.0 kW.
- The machine's power during normal operation (fans) is P_norm ≈ 0.08 kW.

Step-by-step calculation:

- 1. $p = 20 / 180 \approx 0.111$ (11.1 per cent).
- 2. $E(X) = 58 \times 0.111 = 6.4$ aircraft.
- 3. $\sigma = \sqrt{(58 \times 0.111 \times 0.889)} = \sqrt{5.73} \approx 2.4$ aircraft.
- 4. Verification: $n \cdot p \cdot (1-p) = 5.7$ → the normal approximation is suitable for a preliminary estimate, but is at the limits of its validity range, so the result is verified by an exact calculation.
- 5. Confidence level 99.9 per cent, normal approximation: $k = \lceil 6.4 + 3.09 \times 2.4 \rceil = \lceil 13.8 \rceil = 14$ aircraft.
- 6. Exact binomial calculation: $P(X \leq 14) = 99.87$ per cent and $P(X \geq 15) = 0.13$ per cent. The design value k = 14 therefore covers 99.87 per cent of cases; a stricter design value of 99.9 per cent would require 15 aircraft. A value of k = 14 was used in the analysis; the margin of error this leaves (0.03 percentage points) is more than offset by the fact that the underlying assumption of the calculation — that each flat is heated continuously at the highest known heat output — is in itself significantly more stringent than actual usage.
- 7. Concurrency factor: $14/58 \approx 0.24$.
- 8. Scale comparisons based on the exact distribution: $P(X \geq 20) \approx 2 \times 10^{-6}$ (approximately two in a million observations); $P(X = 58) \approx 10^{-55}$.

The binomial distribution in the example case is shown in Figure 1.

Marko Kannisto

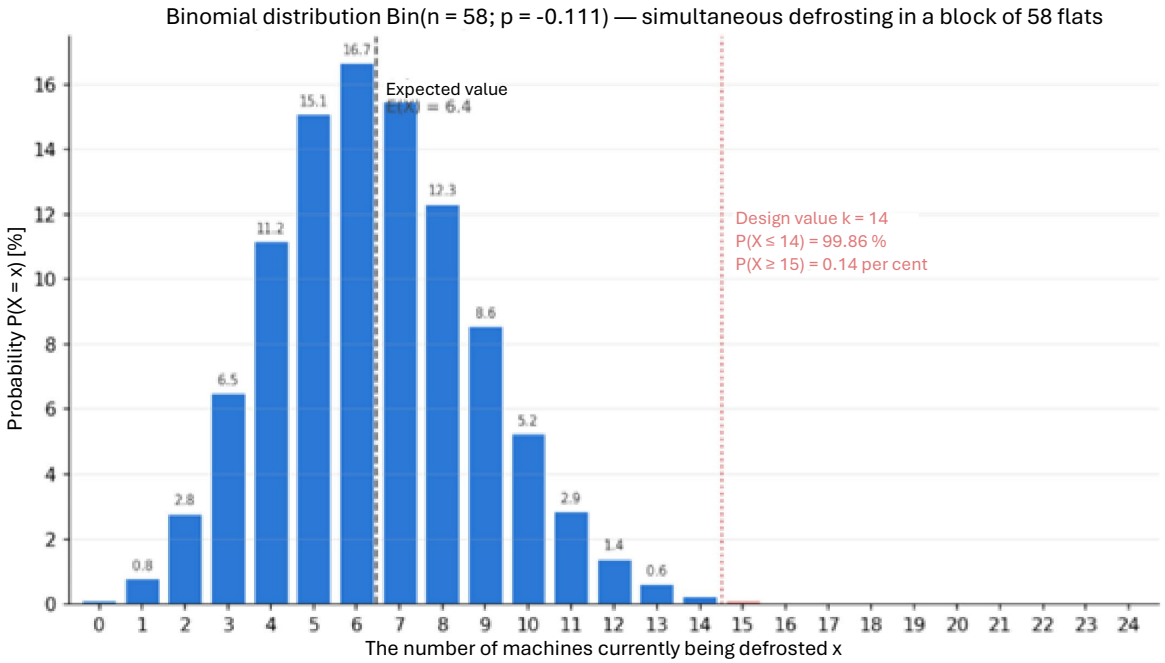

Figure 1. Binomial distribution Bin(n = 58, p = 0.111): the number of IV machines undergoing defrosting at any one time and its probability distribution. The distribution is centred around the expected value (6.4 aircraft) — the most likely single value is 6–7 aircraft. The right-hand tail tapers off rapidly: the design value k = 14 covers 99.86–99.87 per cent of the moments, and the simultaneous defrosting of all 58 machines is, in practice, an impossible scenario.

Rated power:

If the building's electrical system were to be dimensioned on the assumption that all the machines would trip simultaneously, the design load would be 58 × 2.0 kW = 116.0 kW.

In reality, the load consists of two parts: k = 14 machines operating at maximum power during defrosting and the remaining 44 machines operating at normal power. Under normal conditions, heat recovery covers almost all of the heating of the supply air, even in severe frost, in which case the supplementary heating element is hardly ever used, and the unit's electricity consumption is effectively limited to the fans (typically around 80 W). The actual peak load for the building's HVAC systems is calculated as:

14 × 2.0 kW + 44 × 0.08 kW ≈ 31.5 kW

The actual peak load is therefore just over a quarter (around 27 per cent) of the design assumption, even though the calculation takes into account all the building's machinery and a very high level of safety.

The conclusion is not sensitive to the assumption regarding power consumption in normal mode: even if some of the machines operating in normal mode were to use post-heating from time to time and their average power were, for example, 300 W instead of 80 W, the building's peak load would rise to only around 41 kW (14 × 2.0 kW + 44 × 0.3 kW ≈ 41.2 kW) — still well below 40 per cent of the theoretical maximum of 116 kW.

This does not mean that the laws of physics would change or that less heating power would be required. This means that the actual usage profile of the appliances can be taken into account in the sizing process — in the same way that, in electrical design, balancing factors are already applied to the loads of, for

example, saunas and cookers. The binomial distribution provides a transparent, mathematically derived basis for this adjustment, rather than a coefficient based solely on experience.

Exact binomial calculation using recursion (example)

In the example case (n = 58, p = 1/9, where p/(1−p) = 1/8 = 0.125), the recursive calculation proceeds as follows:

x	**The ratio (n−x)/(x+1) × 0.125**	**P(X = x)**	**Cumulative distribution function P(X ≤ x)**
0	—	$(8/9)^{58}$ = 0.0011	0.11%
1	58/1 × 0.125 = 7.250	0.0078	0.89%
2	57/2 × 0.125 = 3.563	0.0279	3.68%
3	56/3 × 0.125 = 2.333	0.0651	10.19%
…	…	…	…
13	46/13 × 0.125 = 0.442	0.0062	99.62%
14	45/14 × 0.125 = 0.402	0.0025	99.87%
15	44/15 × 0.125 = 0.367	0.0009	99.96%

4. Design recommendation

- **Derive the concurrency factor; do not select it.** A multiplier of the ‘utilisation rate + a margin that seems reasonable’ type falls apart at the very first question: ‘What is this based on?’ The binomial distribution produces the same structure (expected value + margin), but the size of the margin is transparently determined by the chosen confidence level.
- **Check the assumption of independence before performing the calculation.** Check the device’s control logic to ensure that the trigger is demand-based. If the cycles are triggered by a common threshold value, time control, the restoration of power following a power cut, or centralised control, the method cannot be applied as it stands.
- **Select the confidence level based on the duration of exposure.** A 95 per cent instantaneous confidence level may sound high, but in practice it is almost certain to be exceeded over a long load cycle. For purposes such as fuse rating, the instantaneous purity must be at least 99.9 per cent.
- **Use the normal approximation only as a rough estimate.** The final design value and the reported confidence level are always taken from an exact binomial calculation, as the approximation may, at the limits of its validity range, yield a value that is one unit too low.
- **Show the initial values in their correct positions.** If the cycle interval is dictated by the device’s control logic, it can be said that ‘the machine cannot operate at a higher frequency than this’. If this is a rule of thumb based on practical experience, the correct wording is ‘typically no more than, based on practical experience’ — otherwise, the document will contain a technical guarantee that cannot be honoured.
- **Specify the entire load right through to the end.** The design value k indicates only the number of peak power loads; please also specify the power at which the remaining n − k loads are calculated. Leaving this open will only lead to further questions.

Pitfalls:

- Load periods of 15–20 minutes are, from the fuse’s point of view, effectively a continuous load, so they cannot be dismissed as a ‘momentary spike’. The value of the method lies in the fact that it reduces the number of loads to be calculated simultaneously — not in the fact that the duration of the load diminishes its significance.
- Independence is a claim that must be substantiated — the mathematics behind the method is trivial, but its applicability depends on an understanding of the triggering mechanism.
- If any input data is missing or inconsistent, this must be clarified before the calculation is carried out — you should not make your own assumptions about the value.

Appendix: Example prompt for AI-assisted calculation

You can also have the calculation carried out by an AI assistant (Claude, Copilot, etc.). The instructions below describe the method in such detail that an assistant will be able to produce the calculation and result structure in accordance with these instructions. Fill in the details in the square brackets and paste the prompt into the conversation as it is.

Two points to note: the language models occasionally make arithmetic errors, which is why the prompt forces the assistant to display intermediate results — please check at least the design value k in Excel before using the result in your design calculations. The guideline also requires the use of an exact binomial calculation rather than a simple normal approximation, as the approximation may, at the boundaries of its validity range, yield a design value that is one unit too low.

Worked as an expert in statistical design calculations. Calculate the simultaneity of periodic electrical loads using the binomial distribution, in accordance with the input data and method description below. Follow the procedure exactly.

INITIAL DATA

- Number of loads: n = [NUMBER, e.g. 58]
- Description of the load: [e.g. the defrosting cycle of a heat recovery unit in a ventilation system]
- Duration of the peak power period: t = [e.g. 20 mins]
- The interval between episodes in the worst-case scenario: T = [e.g. 180 min]
- Load power during the peak power period: P_peak = [e.g. 2.0 kW]
- Load power at other times: P_norm = [value, or "not used in the calculation"]
- Desired instantaneous confidence level: [default 99.9%]
- The episode’s trigger mechanism: [brief description, e.g. "the defrosting process is triggered on a flat-by-flat basis when the coil becomes frosted"]

METHOD - Follow these steps in order and show all intermediate results:

1. Assess the assumption of independence: do the cycles start independently of one another, based on load-specific demand? If the trigger mechanism is shared (time-based control, a threshold value for a shared variable, centralised control), state that the binomial model cannot be applied as it stands, and stop. Please note: a common environmental factor (e.g. outdoor temperature) does not compromise independence when the calculation is carried out conditionally for the worst-case environmental scenario – it determines the frequency of the cycles, not their timing.
2. Calculate the utilisation rate p = t / T.
3. Calculate the expected value E(X) = n·p and the standard deviation σ = √(n·p·(1−p)).

4. Using the normal approximation, calculate an initial estimate k_approx = ⌈E(X) + z·σ⌉, where z is the one-tailed quantile corresponding to the confidence level (99.9% → z = 3.09). Also state n·p·(1−p) and assess the validity of the approximation (range 5–10).
5. Calculate the EXACT binomial distribution using recursion: P(X=0) = (1−p)^n and P(X=x+1) = P(X=x) · (n−x)/(x+1) · p/(1−p). Accumulate the cumulative distribution function P(X ≤ x) and produce a table of the values x, P(X = x) and P(X ≤ x) at least for the range x = 0 to k+2. Define the design value k as the smallest value of x for which the cumulative sum exceeds the confidence level. State the exact coverage P(X ≤ k) and the probability of exceeding k, P(X ≥ k+1), for the selected k.
6. Calculate the simultaneity factor c = k/n and the peak power load k × P_peak. If P_norm is given, also calculate the design power P_design = k × P_peak + (n−k) × P_norm.
7. 7. Calculate the periodic exceedance risk: if a load cycle comprises M time windows of duration t, the exceedance risk during the cycle is ≈ 1 − (1 − α)^M, where α = 1 − P(X ≤ k). Take, for example, a continuous three-day load cycle.
8. Write a summary including: the design value k, its exact coverage, the simultaneity factor, the power results, and the boundary conditions of the calculation (the basis for independence, the basis for the time interval expressed in a form that corresponds to its position – "the device cannot" only if the limit is imposed by the control logic; otherwise, "based on experience typically").

IMPORTANT!

- Never choose the concurrency factor arbitrarily or by applying a 'suitable' 'margin' – the factor is always derived from the distribution.
- The reported level of assurance is derived from a precise calculation, not an approximation.
- If any input data is missing or inconsistent, check it before performing the calculation – do not assume a value yourself.
- Show the calculations so that the reader can check them using Excel (BINOMIAL.DIST.INV and BINOMIAL.DIST).