

PHYSICS FOR ENGINEERING STUDENTS

Probability distributions II

Binomial, Poisson, normal and Cauchy–Lorentz models

Lecture 7 — Book-style student notes

Learning focus: Binomial counts · Poisson events · Normal distribution · Lorentzian line shapes

University of Szeged · Institute of Physics

Probability distributions II

A probability distribution is selected from the physical mechanism and experimental design, not simply from the shape of a histogram. The binomial model describes successes in a fixed number of independent trials; the Poisson model describes independent events occurring with a stable average rate; the normal model describes continuous variation generated by many small contributions; and the Cauchy–Lorentz model describes heavy-tailed phenomena and resonant line shapes.

Learning objectives

- Select binomial and Poisson models for count data.
- Use mean–variance relations and counting-time scaling.
- Standardize normal variables and apply Gaussian approximations.
- Distinguish Gaussian and Cauchy–Lorentz shapes physically.

Central principle: Choose, evaluate, and report a probability or uncertainty model from the physical measurement process and its assumptions.

Further interpretation

Each model comes with assumptions, parameters, and diagnostic implications. The mean–variance relationship, tail behaviour, support, and response to averaging can distinguish distributions that look similar near their centre. Measurement science uses these properties both to predict random uncertainty and to detect when an instrument or physical source violates an assumed model.

1. Binomial experiments

A binomial experiment consists of n predetermined independent Bernoulli trials. Each trial has the same success probability p and failure probability $q=1-p$, and the random variable K counts successes. The binomial coefficient $C(n,k)=n!/[k!(n-k)!]$ counts the distinct sequences containing k successes.

$$P(K = k) = C(n, k)p^k(1 - p)^{n-k}$$

The probability mass function is defined for $k=0, \dots, n$ and sums to one. Changing p moves the centre and changes skewness: $p=0.5$ gives symmetry around $n/2$, small p gives a right-skewed distribution, and p near one gives a left-skewed distribution. Independence can fail when trials share a fluctuating environment, when one event changes later probabilities, or when a detector cannot record closely spaced events.

Further interpretation

A binomial experiment consists of n predetermined independent Bernoulli trials. Each trial has the same success probability p and failure probability $q=1-p$, and the random variable K counts successes. The binomial coefficient $C(n,k)=n!/([k!(n-k)!])$ counts the distinct sequences containing k successes. The probability mass function is defined for $k=0,\dots,n$ and sums to one. Changing p moves the centre and changes skewness: $p=0$. The binomial description is appropriate only when the experimental unit and success event are unambiguous. Repeated trials must be comparable, and dependence between outcomes must be negligible. If the success probability drifts, a mixture of binomial distributions can mimic excess noise and invalidate the nominal uncertainty.

2. Mean, variance and measurement-bin counts

For $K \sim \text{Bin}(n,p)$, $E[K]=np$, $\text{Var}(K)=np(1-p)$, and $\sigma=\sqrt{np(1-p)}$. These results quantify expected successes and their fluctuation. Relative standard deviation $\sigma/E[K]=\sqrt{(1-p)/(np)}$ decreases as n increases, so proportions become more stable in larger independent samples.

$$E[K] = np$$

$$\text{Var}(K) = np(1-p)$$

$$N_j \sim \text{Bin}(N, p_j)$$

Histogram bin counts are themselves random. If N independent measurements are assigned to a particular bin with probability p_j , then its count N_j is binomial: $N_j \sim \text{Bin}(N, p_j)$. Counts in different bins are not independent because their sum equals N . The visible roughness of a finite histogram therefore reflects sampling variability, not necessarily real fine structure in the underlying density.

Further interpretation

For $K \sim \text{Bin}(n,p)$, $E[K]=np$, $\text{Var}(K)=np(1-p)$, and $\sigma=\sqrt{np(1-p)}$. These results quantify expected successes and their fluctuation. Relative standard deviation $\sigma/E[K]=\sqrt{(1-p)/(np)}$ decreases as n increases, so proportions become more stable in larger independent samples. Histogram bin counts are themselves random. If N independent measurements are assigned to a particular bin with probability p_j , then its count N_j is binomial: $N_j \sim \text{Bin}(N, p_j)$. Counts in different bins are not independent because their sum equals N . The mean predicts the expected bin occupancy, while the variance predicts repeat-to-repeat fluctuation. Empty or saturated bins carry information as well as populated ones. Comparing observed count variation with $Np(1-p)$ helps reveal changing probabilities, correlations, or an incorrectly defined set of trials.

3. Counting random events and the Poisson distribution

The Poisson model applies to the number of independent events occurring in a fixed interval of time, space, area, or volume when the average rate is stable and simultaneous events have negligible probability. Examples include photon arrivals, radioactive decays, cosmic-ray detections, defects per area, and failures over time.

$$P(K = k) = e^{-\lambda} \lambda^k / k!$$

$$\lambda = r\Delta t$$

If λ is the expected number of events in the interval, then $P(K=k)=e^{-\lambda} \lambda^k / k!$ for $k=0,1,\dots$. The parameter is dimensionless because it is an expected count. For a constant event rate r observed for time Δt , $\lambda=r\Delta t$. Doubling the observation interval doubles the expected count, provided the source remains stationary and detector losses are negligible.

Further interpretation

The Poisson model applies to the number of independent events occurring in a fixed interval of time, space, area, or volume when the average rate is stable and simultaneous events have negligible probability. Examples include photon arrivals, radioactive decays, cosmic-ray detections, defects per area, and failures over time. If λ is the expected number of events in the interval, then $P(K=k)=e^{-\lambda} \lambda^k / k!$ for $k=0,1,\dots$. The Poisson model describes independent events occurring at a stable average rate within a specified exposure. Both the event definition and the observation interval matter. Dead time, pile-up, bursts, and rate drift violate the assumptions and often produce variance that differs from the mean.

4. Mean, variance and shape of Poisson counts

A Poisson random variable has mean and variance both equal to λ , so its standard deviation is $\sqrt{\lambda}$. For small λ the distribution is strongly right-skewed and zero events may be most probable. As λ increases, the distribution becomes more symmetric and increasingly resembles a Gaussian centred at λ .

$$E[K] = \text{Var}(K) = \lambda$$

$$\sigma_K = \sqrt{\lambda}$$

$$\gamma_1 = 1/\sqrt{\lambda}$$

$$\gamma_{2,\text{excess}} = 1/\lambda$$

Poisson skewness is $\frac{1}{\sqrt{\lambda}}$ and excess kurtosis is $1/\lambda$. These decrease with increasing expected count. Observed variance substantially larger than the mean indicates overdispersion caused by changing

rate, clustering, background variation, or heterogeneous sources. Variance smaller than the mean can occur through dead time, exclusion, conservation constraints, or other negative dependence.

Further interpretation

A Poisson random variable has mean and variance both equal to λ , so its standard deviation is $\sqrt{\lambda}$. For small λ the distribution is strongly right-skewed and zero events may be most probable. As λ increases, the distribution becomes more symmetric and increasingly resembles a Gaussian centred at λ . Poisson skewness is $1/\sqrt{\lambda}$ and excess kurtosis is $1/\lambda$. These decrease with increasing expected count. Relative Poisson noise decreases as one over the square root of the expected count, but absolute fluctuation increases. Longer counting improves relative precision only while the rate and background remain stable. This trade-off connects acquisition time directly to the uncertainty target of the experiment.

5. From binomial to Poisson

The Poisson distribution is obtained as a limit of the binomial distribution when n becomes large, p becomes small, and $np=\lambda$ remains finite. Physically, there are many opportunities for a rare event, while the expected total count stays moderate. Under these conditions $C(n,k)p^k(1-p)^{n-k}$ approaches $e^{-\lambda} \lambda^k/k!$ for each fixed k .

$$n \rightarrow \infty, p \rightarrow 0, np = \lambda$$

$$\text{Bin}(n, p) \rightarrow \text{Poisson}(\lambda)$$

Use the binomial model when the number of trials n is fixed, and success probability p is defined for each trial. Use Poisson when events occur in an interval and only the average rate is specified. For $n=100$ and $p=0.01$, $\lambda=1$, and the Poisson approximation is close, although the binomial variance $np(1-p)=0.99$ remains slightly below the Poisson variance 1.

Further interpretation

The Poisson distribution is obtained as a limit of the binomial distribution when n becomes large, p becomes small, and $np=\lambda$ remains finite. Physically, there are many opportunities for a rare event, while the expected total count stays moderate. Under these conditions $C(n,k)p^k(1-p)^{n-k}$ approaches $e^{-\lambda} \lambda^k/k!$ for each fixed k . Use the binomial model when the number of trials n is fixed, and success probability p is defined for each trial. The Poisson approximation to a binomial model is strongest for many opportunities and small success probability while Np remains finite. Its convenience does not remove the need to check bounds: a binomial count cannot exceed N , whereas a Poisson model assigns small probability beyond that limit.

6. Poisson additivity and thinning

Independent Poisson counts add: if $K_1 \sim \text{Poisson}(\lambda_1)$ and $K_2 \sim \text{Poisson}(\lambda_2)$, then $K_1 + K_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$. This represents combining independent sources or adjacent observation intervals. The property supports background-plus-signal models and the accumulation of counts over time.

$$K_1 + K_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$$

$$K_{\text{detected}} \sim \text{Poisson}(\eta\lambda)$$

Thinning preserves the Poisson form. If every event is detected independently with efficiency η , the detected count is Poisson with mean $\eta\lambda$, while missed events form an independent Poisson count with mean $(1-\eta)\lambda$. A fluctuating efficiency or dead time violates simple thinning. These properties explain why stable losses change the mean but do not by themselves create overdispersion.

Further interpretation

Independent Poisson counts add: if $K_1 \sim \text{Poisson}(\lambda_1)$ and $K_2 \sim \text{Poisson}(\lambda_2)$, then $K_1 + K_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$. This represents combining independent sources or adjacent observation intervals. The property supports background-plus-signal models and the accumulation of counts over time. Thinning preserves the Poisson form. If every event is detected independently with efficiency η , the detected count is Poisson with mean $\eta\lambda$, while missed events form an independent Poisson count with mean $(1-\eta)\lambda$. Independent Poisson streams add because their rates add. Thinning preserves the Poisson form when each event is accepted independently with fixed probability. These properties explain detector efficiency corrections, channel splitting, and combined backgrounds, but they fail when acceptance depends on event history or intensity.

7. Photon counting and shot noise

For ideal independent photon arrivals, the count N during Δt is Poisson with mean $\lambda = r\Delta t$ and standard deviation $\sqrt{\lambda}$. The relative standard deviation is $1/\sqrt{\lambda}$. Because expected count grows linearly with observation time while noise grows only with its square root, longer counting improves relative precision as $1/\sqrt{\Delta t}$.

$$u_r(N) = \sigma_N / E[N] = 1/\sqrt{\lambda}$$

$$\lambda = r\Delta t$$

This improvement assumes constant signal and background. Dark counts and background photons add to the observed mean and noise; subtracting an independently measured background adds its variance. Detector dead time suppresses closely spaced counts, afterpulsing creates correlated extra pulses, and source intensity fluctuation produces excess noise. Agreement of count variance with mean is therefore a useful diagnostic of ideal shot-noise-limited operation.

Further interpretation

For ideal independent photon arrivals, the count N during Δt is Poisson with mean $\lambda = r\Delta t$ and standard deviation $\sqrt{\lambda}$. The relative standard deviation is $1/\sqrt{\lambda}$. Because expected count grows linearly with observation time while noise grows only with its square root, longer counting improves relative precision as $1/\sqrt{\Delta t}$. This improvement assumes constant signal and background. Shot noise is a consequence of discrete arrival events and therefore scales with the square root of the detected count. Improving optical power or integration time raises signal faster than shot-noise standard deviation, but electronic noise, dark counts, saturation, and drift can dominate outside the shot-noise-limited regime.

8. The normal distribution

A continuous variable X is normally distributed with mean μ and standard deviation σ when its density is $f(x) = [1/(\sigma\sqrt{2\pi})]\exp[-(x-\mu)^2/(2\sigma^2)]$. The curve is symmetric, bell shaped, and normalized. Mean, median, and mode coincide at μ ; changing μ shifts the curve, while increasing σ makes it wider and lower without changing total area.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

The normal distribution has zero skewness and zero excess kurtosis. Approximately 68.27% of probability lies within $\mu \pm \sigma$, 95.45% within $\mu \pm 2\sigma$, and 99.73% within $\mu \pm 3\sigma$. These percentages are properties of the model, not universal rules for arbitrary measurement data. Thresholds, saturation, multiplicative errors, mixtures, and heavy tails can make them invalid.

Further interpretation

A continuous variable X is normally distributed with mean μ and standard deviation σ when its density is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. The curve is symmetric, bell shaped, and normalized. Mean, median, and mode coincide at μ ; changing μ shifts the curve, while increasing σ makes it wider and lower without changing total area. The normal distribution has zero skewness and zero excess kurtosis. Approximately 68.27% of probability lies within $\mu \pm \sigma$, 95. A normal model is fully determined by its mean and standard deviation, but those parameters do not guarantee normality. Symmetry, light tails, and a plausible additive mechanism should also be checked. Standardized residual plots are often more revealing than a histogram alone for limited samples.

9. Standardization and interval probability

Any normal variable can be standardized using $Z=(X-\mu)/\sigma$. The resulting variable follows the standard normal distribution with mean zero and standard deviation one. Probability calculations for different μ and σ then use one universal CDF $\Phi(z)$: $P(a \leq X \leq b) = \Phi((b-\mu)/\sigma) - \Phi((a-\mu)/\sigma)$.

$$Z = (X - \mu)/\sigma$$

$$P(a \leq X \leq b) = \Phi((b - \mu)/\sigma) - \Phi((a - \mu)/\sigma)$$

A z-score states how many standard deviations an observation lies from the mean. It is dimensionless and preserves order, but a large $|z|$ should be interpreted in context. When μ and σ are estimated from the same small sample, Studentized quantities may follow a t distribution rather than an exact standard normal distribution.

Further interpretation

Any normal variable can be standardized using $Z=(X-\mu)/\sigma$. The resulting variable follows the standard normal distribution with mean zero and standard deviation one. Probability calculations for different μ and σ then use one universal CDF $\Phi(z)$: $P(a \leq X \leq b) = \Phi((b-\mu)/\sigma) - \Phi((a-\mu)/\sigma)$. A z score states how many standard deviations an observation lies from the mean. It is dimensionless and preserves order, but a large $|z|$ should be interpreted in context. Standardization converts a value to its displacement from the mean in standard-deviation units. It supports use of common probability tables and comparison across scales. The resulting probability remains conditional on the normal model and on correctly estimated parameters; it is not a distribution-free confidence statement.

10. Why Gaussian behavior is common

Many measured quantities are influenced by sums of numerous small contributions: electronic noise sources, positioning variations, thermal fluctuations, material inhomogeneity, and reading effects. The central limit theorem states that, under broad conditions of independence and finite variance, a suitably standardized sum tends toward a normal distribution as the number of contributions grows.

$$S_n = Y_1 + Y_2 + \dots + Y_n$$

The theorem concerns sums and averages, not every physical variable. A product of positive random factors may be closer to lognormal; a quantity dominated by one rare event may be skewed; and heavy-tailed contributions with infinite variance need not converge to a Gaussian. Dependence can also slow or prevent convergence. Normality should be assessed with physical reasoning and diagnostic plots rather than assumed from tradition.

Further interpretation

Many measured quantities are influenced by sums of numerous small contributions: electronic noise sources, positioning variations, thermal fluctuations, material inhomogeneity, and reading effects.

The central limit theorem states that, under broad conditions of independence and finite variance, a suitably standardized sum tends toward a normal distribution as the number of contributions grows. The theorem concerns sums and averages, not every physical variable. Gaussian behavior often emerges when many small, weakly dependent contributions add and no single term dominates. The argument is about sums or averages, not necessarily individual physical effects. Strong dependence, heavy tails, thresholds, or mixtures can prevent the approximation from becoming accurate.

11. Gaussian approximations to counts

For large n , when both np and $n(1-p)$ are sufficiently large, a binomial distribution can be approximated by a normal distribution with $\mu=np$ and $\sigma^2=np(1-p)$. A continuity correction improves probability estimates because the discrete count k represents an interval roughly from $k-0.5$ to $k+0.5$ on the continuous scale.

$$\text{Bin}(n, p) \approx N(np, np(1 - p))$$

$$\text{Poisson}(\lambda) \approx N(\lambda, \lambda)$$

A Poisson distribution with large λ is approximated by $N(\lambda, \lambda)$. The approximation is least reliable in the far tails and near the boundary at zero. Exact binomial or Poisson calculations are preferred when feasible, particularly for small counts, rare-event decisions, and confidence limits where asymmetry matters. Convenience should not override a model whose limitations affect the conclusion.

Further interpretation

For large n , when both np and $n(1-p)$ are sufficiently large, a binomial distribution can be approximated by a normal distribution with $\mu=np$ and $\sigma^2=np(1-p)$. A continuity correction improves probability estimates because the discrete count k represents an interval roughly from $k-0.5$ to $k+0.5$ on the continuous scale. A Poisson distribution with large λ is approximated by $N(\lambda, \lambda)$. A Gaussian approximation to counts should be checked against expected count size and tail location. Continuity correction improves interval estimates by respecting discrete bin boundaries. For rare-event tails, direct binomial or Poisson calculation is usually safer because small absolute errors can become large relative errors.

12. The Cauchy–Lorentz distribution and line shapes

The Cauchy–Lorentz density centred at x_0 with half-width γ is $f(x) = \gamma / \{\pi[(x - x_0)^2 + \gamma^2]\}$. Its FWHM is 2γ . It has a sharp central peak and tails that decrease as $1/x^2$, much more slowly than Gaussian tails. The usual mean and variance integrals do not converge, so averaging does not acquire the familiar Gaussian stability.

$$f(x) = \gamma / \{\pi[(x - x_0)^2 + \gamma^2]\}$$

$$\text{FWHM} = 2\gamma$$

Lorentzian profiles arise naturally from exponentially damped oscillations and resonance phenomena. The Fourier transform of $Ae^{-\gamma t} \cos(\omega_0 t + \phi)$ produces a spectral feature centred near ω_0 with width related to the damping rate. Instrumental Gaussian broadening combined with physical Lorentzian broadening produces a Voigt profile. Selecting Gaussian or Lorentzian width therefore carries physical information about broadening mechanisms, not merely curve-fitting preference.

Take-home message: The mathematical expression is useful only together with the physical assumptions that justify it and the uncertainty statement that limits its interpretation.

Further interpretation

The Cauchy–Lorentz density centered at x_0 with half-width γ is $f(x) = \gamma / \{\pi[(x - x_0)^2 + \gamma^2]\}$. Its FWHM is 2γ . It has a sharp central peak and tails that decrease as $1/x^2$, much more slowly than Gaussian tails. The usual mean and variance integrals do not converge, so averaging does not acquire the familiar Gaussian stability. Lorentzian profiles arise naturally from exponentially damped oscillations and resonance phenomena. Lorentzian profiles have much heavier tails than Gaussian profiles and do not possess a finite variance. Consequently, Gaussian width-to-standard-deviation conversions are invalid. Model choice should be based on the physical broadening mechanism and residuals, especially when fitted peak area or far-wing behavior affects the conclusion.

13. Integrated measurement workflow and worked applications

The following compact studies connect the definitions and equations to realistic interpretation, calculation, and experimental design decisions. A practical model-selection workflow starts by identifying how counts or continuous readings are produced. A fixed number of comparable two-outcome trials suggests a binomial model; independent arrivals during a fixed exposure suggest Poisson statistics; a sum of many small contributions may justify a normal approximation; and a resonance or lifetime-broadened line may require a Lorentzian profile. The support and generating mechanism are stronger evidence than superficial visual similarity. After choosing a candidate, estimate its parameters and compare predicted and observed variability. Binomial data should be

checked against $Np(1-p)$, Poisson counts against equality of mean and variance, and normal residuals against symmetry and tail behavior. Time plots are essential because drift, bursts, dead time, and changing efficiency can be hidden by an acceptable overall histogram. Background and efficiency corrections should retain their own counting uncertainty and any shared components. Approximations should be documented as approximations. When replacing a discrete count model by a Gaussian, check the expected counts, apply a continuity correction for intervals, and avoid relying on the approximation in rare-event tails. When fitting peaks, compare residuals and tail behavior before converting widths between Gaussian and Lorentzian conventions. The final report should name the distribution, parameter estimates, exposure or trial definition, correction model, and diagnostic evidence. This makes the calculated probability traceable to the actual measurement process.

Binomial quality control

If each component independently fails with probability $p=0.02$ and 50 components are tested, the number of failures is binomial with mean 1 and variance 0.98. The probability of no failures is $0.98^{50} \approx 0.364$. Observing zero failures is therefore not strong evidence that the failure probability is zero. A fixed sample size and binary outcome justify the binomial model; if failure probability varies across components or production batches, the observed distribution becomes more dispersed than a single- p binomial model.

Poisson photon counts

A detector receives an average of $\lambda=4$ photons per gate. The probabilities are $P(K=k)=e^{-4}4^k/k!$. The mean and variance are both four, so the standard deviation is two and the relative standard deviation is 50%. A single count of two does not imply that the source intensity halved; it is a plausible Poisson fluctuation. Averaging many gates estimates the rate more precisely, while the gate-to-gate variance provides a diagnostic of whether the source is stable and shot-noise limited.

Counting time trade-off

A stable source produces $r=100$ counts s^{-1} . In 1 s, $\lambda=100$ and ideal relative counting noise is 10%. In 100 s, $\lambda=10,000$ and it falls to 1%. The improvement follows $1/\sqrt{\Delta t}$, so a tenfold reduction requires a hundredfold longer measurement. Very long integration may become counterproductive if the source drifts, background changes, or the measurand itself evolves. Counting time should therefore be chosen by balancing statistical precision against temporal resolution and stability.

Background subtraction

A signal-plus-background measurement gives 150 counts, and an independent equal-time background measurement gives 50 counts. The net estimate is 100 counts. For independent Poisson variables, variances add under subtraction, so the net variance is $150+50=200$ and the standard deviation is about 14.1, not $\sqrt{100}=10$. Background subtraction removes the expected background level but not its

counting noise. A longer background measurement can reduce its rate uncertainty after appropriate scaling.

Poisson thinning

An ideal source produces Poisson counts with $\lambda=100$, but a detector independently records each event with efficiency $\eta=0.30$. The detected count is Poisson with mean 30 and standard deviation $\sqrt{30}\approx 5.48$. Correcting the mean by division by η gives an unbiased source-count estimate, but its uncertainty is also divided by η . If η has calibration uncertainty, that Type B contribution must be propagated. At high rates, dead time makes detection events dependent, and the thinning result no longer applies.

Normal interval probability

For $X\sim N(50,3^2)$, the interval $47\leq X\leq 56$ corresponds to z values -1 and $+2$. Its probability is $\Phi(2)-\Phi(-1)\approx 0.9772-0.1587=0.8185$. Standardization separates location and scale from the universal normal shape. The same z interval has the same probability for any μ and σ . When parameters are estimated from a small sample, however, uncertainty in μ and σ must be considered rather than treating them as perfectly known constants.

Continuity correction

Let $K\sim \text{Bin}(100,0.5)$. To approximate $P(K\leq 55)$ with a normal distribution, use $\mu=50$, $\sigma=5$ and replace the discrete boundary 55 by 55.5. The standardized value is 1.1, giving approximately $\Phi(1.1)=0.864$. Without continuity correction, $z=1.0$ gives 0.841 and a noticeably poorer approximation. The correction represents each integer count by a unit-width interval. Exact binomial calculation remains preferred when software is available, and tail accuracy matters.

Gaussian and Lorentzian peaks

Two peaks can have the same center and FWHM but very different tails. A Gaussian decays exponentially in squared distance and has finite variance; a Lorentzian decays as $1/(x-x_0)^2$ and has undefined variance. Far from the line center, Lorentzian intensity is much larger. Residuals on a logarithmic vertical scale reveal this difference better than a linear plot dominated by the peak. In spectroscopy, Doppler broadening tends to be Gaussian, lifetime or collision broadening Lorentzian, and their convolution a Voigt profile.

Integrated interpretation

The distributions in this lecture form a connected hierarchy of measurement models. The binomial model begins with a fixed number of comparable opportunities and a success probability. The Poisson model emerges for rare independent events in a fixed exposure and introduces the characteristic equality of mean and variance. At sufficiently large counts, both may be approximated by a normal curve near their centres, but their discrete support and asymmetric tails remain important for rare-

event decisions. Gaussian and Lorentzian continuous profiles can share the same center and FWHM while representing radically different tail behavior, so width alone cannot select the physical model. A complete analysis should therefore proceed from mechanism to assumptions, parameters, diagnostics, and only then to probability. Define the trial or event, exposure, efficiency, background, and any dead-time correction. Compare observed fluctuation with the variance predicted by the candidate distribution and inspect the sequence for drift or clustering. Use exact count probabilities when expected numbers are small or tails matter; document continuity correction when a normal approximation is used. For peaks, examine residuals over both the center and wings and identify the broadening mechanism before interpreting fitted widths or areas. This sequence also clarifies experimental design. Increasing trials, exposure time, or detected intensity reduces relative counting uncertainty only while the process remains stationary and the detector remains linear. Background measurements must receive enough exposure to avoid dominating a subtraction. Efficiency improvements change both expected signal and its counting variance. The final result should state the distribution, parameter estimates, conditions, corrections, and evidence supporting adequacy. Probability is then not merely a collection of formulas but a testable bridge between microscopic event generation and the uncertainty of a macroscopic measurement result.