Poisson model

For CMB fluctuations it's easy to know how to relate theory, $C(\theta)$, to observations at $T(\hat{n})$.
How do we connect

$$\xi(\vec{r}) \equiv \langle \delta(\vec{x}+\vec{r})\,\delta(\vec{x}) \qquad \text{vs.} \qquad \delta P_{12} = \bar{n}\,\delta V\;\bar{n}\,\delta V\,(1+\xi_{12})$$

for galaxies?

Introduce a model to allow us to go from cts fields to point processes. (discrete objects)
This model adds stochasticity, makes assumptions and is only approximate.

Poisson model (Layzer, Limber)

First select $\rho(\vec{r})$ from ensemble
Place an object with probability $\delta P = \rho(\vec{r})\,\delta V$, indep. of other objects.

Note This can't be whole story - deviations due to volume exclusion (hard spheres) or excess clustering.
A more clever stat. formulation can be found in "Theory of point processes" (Daley & Vere-Jones)
Non-poisson from N-body see Casas-Miranda et al. [a-p/0105008]

Note that this model does what we want

$$\delta P = \rho(r_1)\,\delta V_1\,\rho(r_2)\,\delta V_2$$

average over ρ's $\quad \langle \rho(x)\rho(x+r)\rangle = \bar{\rho}^2(1+\xi(r))$ so $\quad \delta P = \bar{\rho}^2\,\delta V_1\,\delta V_2\,[1+\xi(r)]$

Now when we compute 2-point fns there are 2 levels of averaging

$$\langle n_i n_j \rangle \equiv \langle\langle \bar{n}(1+\delta_i)\bar{n}(1+\delta_j) \rangle_c \rangle_p \qquad \text{for counts in cells } i \neq j$$
$$= \langle \bar{n}(1+\delta_i)\bar{n}(1+\delta_j) \rangle_c + \delta_{ij} \langle \bar{n}(1+\delta_i) \rangle_c$$
$$= \bar{n}^2 (1+\xi_{ij}) + \bar{n}\delta_{ij}$$

Let's look at this 1st step again. Can use

$$P(n;\mu) = \frac{\mu^n}{n!} e^{-\mu}$$
$$\langle 1 \rangle = 1$$
$$\langle n \rangle = \mu$$
$$\langle n(n-1) \rangle = \mu^2 \Rightarrow \langle n^2 \rangle = \mu + \mu^2 \qquad \mathrm{Var}[n] = \mu = E[n]$$
$$\langle n(n-1)(n-2) \rangle = \mu^3 \Rightarrow \langle n^3 \rangle = \mu + 3\mu^2 + \mu^3$$
etc

along with independence: $\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle$ if $i \neq j$.

But there is an easier way to do it.
Divide space into infinitesimal cells with $\mu \ll 1$. Then $n_i = 0$ or 1 so
$$\langle n \rangle = \langle n^2 \rangle = \langle n^3 \rangle = \cdots$$

So for example

$$|\delta_k|^2 = \int d^3r_1\, d^3r_2\, e^{i\vec{k}\cdot(\vec{r}_1-\vec{r}_2)} \left\langle\left\langle \frac{n_1-\bar{n}}{\bar{n}} \frac{n_2-\bar{n}}{\bar{n}} \right\rangle\right\rangle$$
$$= \sum_{ij} e^{i\vec{k}\cdot(\vec{r}_i-\vec{r}_j)} \left\langle\left\langle \frac{N_i-\bar{N}}{\bar{n}} \frac{N_j-\bar{N}}{\bar{n}} \right\rangle\right\rangle \qquad N_i \equiv n_i\,\delta V = 0,1$$
$$= \sum_{ij} \frac{e^{i\vec{k}(\vec{r}_i-\vec{r}_j)}}{\bar{n}^2} \langle \delta\mu_i\,\delta\mu_j + \delta_{ij}\,\mu_i \rangle$$
$$= \int d^3r_1\, d^3r_2\, e^{i\vec{k}(\vec{r}_1-\vec{r}_2)} \left\{ \left\langle \frac{\delta\rho}{\rho}(r_1) \frac{\delta\rho}{\rho}(r_2) \right\rangle + \delta(r_1-r_2)\frac{1}{\bar{\rho}} \right\}$$

$$= P(k) + \frac{1}{\bar{\rho}}$$

Poisson model

Note we have two indep. causes of variance; power spectra add and white/Poisson noise has $P(k) = \text{const}$.

So our estimate of power would be

$$\hat{P}(k) = |\delta_k|^2 - \frac{1}{\bar{n}}$$

How would we compute δP?
Assume linear theory with δ_k independent Gaussians.
In the limit $\bar{n} \to \infty$

$$\begin{aligned} \mathrm{Var}[\hat{P}] &= \langle \hat{P}^2 \rangle - \langle \hat{P} \rangle^2 \\ &= \langle \delta^4 \rangle - \langle \delta^2 \rangle^2 \\ &= 2P(k) \end{aligned}$$

Note $\langle x^4 \rangle = 3\sigma^4$ for Gaussian

or

$$\frac{\delta P}{P} = \sqrt{\frac{2}{N}} \quad \text{for } N \text{ independent modes.}$$

Now reinstate $\bar{n} < \infty$. If Poisson "noise" is independent of signal

$$\langle (S+N)(S+N) \rangle - \langle S+N \rangle^2 = \langle S^2 \rangle - \langle S \rangle^2 + \langle N^2 \rangle - \langle N \rangle^2 \qquad \text{if } \langle SN \rangle = 0$$

The "S" terms give $2P(k)$ while the "N" terms give $2\frac{1}{\bar{n}}$ so

$$\frac{\delta P}{P} = \sqrt{\frac{2}{N}} \left[1 + \frac{1}{\bar{n}P} \right] \qquad \text{to leading order in } \frac{1}{\bar{n}V}$$

(full result in Meiksin + White MNRAS '99).

Another way of looking at this result: what we measure is random with power $P(k) + \frac{1}{\bar{n}}$.
Subtracting a constant from $\hat{P}$ doesn't change statistics so

$$\frac{\delta\hat{P}}{\hat{P}} = \frac{\delta P}{\hat{P}} = \sqrt{2}\left[1 + \frac{1}{\bar{n}P}\right]$$

<u>Note</u> Have "sample variance" and noise contributions to δP.
First is multiplicative and pushes to maximize sky area (number of modes) while shot-noise is "additive".

Once $\bar{n}P > 1$ reach diminishing returns for δP. At $\bar{n}P = 5$ are 80% to ∞.
Near peak of $P(k)$ have $P \simeq 2500 - 5000\ (h^{-1}\mathrm{Mpc})^3$ so
optimal $\bar{n} \gtrsim 10^{-4}$, at or less than L_* galaxies.

Kaiser - subsample and cover lots of area.
<u>BUT</u> must subsample randomly!!