

Problem set 7

The first three problems are calculations, which are important to do. The rest are exercises are on combinatorial problems that we have considered before. Exact counting arguments can be difficult, but calculating mean and variance is often easy and gives much information about the random variable under consideration.

Problem 1. Find the expectation, variance and moment generating function for the following distributions. (1) $\text{Bin}(n, p)$, (2) $\text{Geo}(p)$, (3) $\text{Poi}(\lambda)$.

Problem 2. Find the expectation, variance and moment generating function for the following distributions. (1) $\text{N}(\mu, \sigma^2)$, (2) $\Gamma(\alpha, \lambda)$, (3) $\text{Beta}(p, q)$. [Note: For the last one, at least do the special case of $\text{Unif}[0, 1]$.]

Problem 3. Show that the following distributions do not have expectation: 1. Cauchy distribution having density $\frac{1}{\pi(1+x^2)}$ on the whole line. 2. Pareto distribution with density $\frac{\alpha}{x^{\alpha+1}}$ for $x > 1$, when $\alpha \leq 1$.

Problem 4. Suppose r distinguishable balls are placed at random into m labelled bins. Find the expectation and variance of the number of empty bins. [Hint: If X is the number of empty bins, the pmf of X is rather complicated. Don't try to do it by direct calculation which is fairly impossible!]

Problem 5. A box contains N coupons labelled $1, 2, \dots, N$. Two coupons are drawn from the box and the numbers added. Find the expectation and variance in both the following scenarios. (1) The coupons are drawn with replacement. (2) The coupons are drawn without replacement.

Problem 6. In the same setting as the previous example, suppose m coupons are drawn at random (consider both cases, with or without replacement). Let X be the total of the coupons drawn. Find $\mathbf{E}[X]$ and $\text{Var}(X)$. [Note: If you solved the previous problem, you can use it to get the answer fairly easily.]

Problem 7. (*) Consider the coupon collector problem: From a box containing N labelled coupons, draws are made with replacement till all the coupons are seen. Let X_N be the total number of draws.

1. Argue (heuristic argument is good enough) that X_N has the same distribution as $X_1 + \dots + X_N$ where X_k s are independent random variables and X_k has $\text{Geo}((N - k + 1)/N)$ distribution.
2. Use the first part to calculate the mean and variance of X .
3. Show that X is close to $N \log N$ with high probability. More precisely, show that $\mathbf{P}\{X_N \in [N \log N - Nh_N, N \log N + Nh_N]\} \rightarrow 1$ for any sequence h_N that goes to infinity as $N \rightarrow \infty$.