Math 340 HW 6

Asa Royal (ajr74) [collaborators: none]

March 8, 2024

1. **Meester 2.7.29**. Suppose that a given experiment has k possible outcomes, the ith outcome having probability p_i. Denote the number of occurrences of the ith outcome having probability p_i. Denote the number of occurences of the ith outcome in n independent experiments by N_i. Show that

$$\mathbb{P}(N_i = n_1, \dots N_k = n_k) = \frac{n!}{n_1!n_2!\dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

Proof. Since the experiments are independent, $\mathbb{P}(N_i = k) = {p_i}^k$ and the chance of seeing a specific ordering of outcomes where $N_1 = 1, N_2 = 2, \dots N_k = k$ is $\prod_{i=1}^{k} {p_i}^{n_k}$. But we are asked for an order-independent probability, so we must multiply that quantity by the total number of ways this could happen. That number is given in the counting writeup on canvas: the number of ways to partition n distinct results of an experiment to k possible outcomes is $n!/(n_1!n_2!\dots n_k!)$. Thus,

$$\mathbb{P}(N_i = n_1, \dots N_k = n_k) = \frac{n!}{n_1!n_2!\dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

□

2. Let $p \in (0, 1)$. Suppose the random variables $X_1, X_2, \dots, X_n$ are independent adn ahve the Bernoulli(p)) distribution.

(i) Compute the function

$$\mu(\lambda) = \ln \mathbb{E}[e^{\lambda X_1}], \lambda \in \mathbb{R}$$

in terms of p and λ. In particular, what are $\mu(0)$ and $\mu'(0)$?

Compute $\mu(\lambda)$:
Per the definition of expected value,

$$\mu(\lambda) = \ln \sum_{x \in R(X)} e^{\lambda x} \mathbb{P}(X_1 = x) \tag{1}$$

Since X_1 is a Bernoulli random variable, it can only take on the values 0 and 1. Thus (1) is equivalent to

$$\begin{aligned} \mu(\lambda) &= \ln \left(\sum_{x=0}^{1} e^{\lambda x} \mathbb{P}(X_1 = x) \right) \\ &= \ln \left(e^{0 * \lambda} * \mathbb{P}(X_1 = 0) + e^{1 * \lambda} * \mathbb{P}(X_1 = 1) \right) \\ &= \boxed{\ln(1 - p + pe^{\lambda})} \end{aligned}$$

Evaluate $\mu(0), \mu'(0)$:
Plugging $\lambda = 0$ into $\mu(\lambda)$, we see

$$\mu(0) = \ln(1 - p + pe^0) = \ln(1) = \boxed{0}$$

Taking the derivative of $\mu(\lambda)$, we find

$$\mu'(\lambda) = \frac{pe^{\lambda}}{e^{\lambda}p + (1 - p)}$$

So

$$\mu'(0) = \frac{pe^0}{e^0 p + 1 - p} = \boxed{p}$$

(ii) Compute the function

$$h(\lambda) = \ln \mathbb{E}\left[e^{\lambda(X_1+\ldots+X_n)}\right]$$

in terms of $\mu(\lambda)$ and n.

$$\begin{aligned}
h(\lambda) &= \ln \mathbb{E}\left[e^{\lambda(X_1+\ldots+X_n)}\right] \\
&= \ln \mathbb{E}\left[e^{\lambda X_1} e^{\lambda X_2} \ldots e^{\lambda X_n}\right] && \text{split exponential term} \\
&= \ln \left(\mathbb{E}\left[e^{\lambda X_1}\right] \mathbb{E}\left[e^{\lambda X_2}\right] \ldots \mathbb{E}\left[e^{\lambda X_n}\right]\right) && \text{because the } X_j \text{ are indep.} \\
&= \ln \mathbb{E}\left[e^{\lambda X_1}\right] + \ln \mathbb{E}\left[e^{\lambda X_2}\right] + \ldots + \ln \mathbb{E}\left[e^{\lambda X_n}\right] && \text{log of product} \to \text{sum of logs} \\
&= n \ln \mathbb{E}\left[e^{\lambda X_1}\right] && \text{because the } X_j \text{ have same dist.} \\
&= n\mu(\lambda)
\end{aligned}$$

(iii) Show that for any $\varepsilon > 0$ and $\lambda > 0$,

$$\mathbb{P}(X_1 + \ldots + X_n > n(p+\varepsilon)) \leq e^{n(\mu(\lambda)-\lambda(\varepsilon+p))} \tag{2}$$

Proof. Consider the function $f(x) = e^{\lambda x}$. Because f increases monotonically with x, for a positive r.v. Y,

$$\mathbb{P}(Y > a) = \mathbb{P}(f(Y) > f(a))$$

So to prove (2), we can show that

$$\mathbb{P}\left(e^{\lambda(X_1+\ldots+X_n)} \geq e^{\lambda n(p+\varepsilon)}\right) \leq e^{n(\mu(\lambda)-\lambda(p+\varepsilon))} \tag{3}$$

Per Markov's inequality,

$$\begin{aligned}
\mathbb{P}\left(e^{\lambda(X_1+\ldots+X_n)} \geq e^{\lambda n(p+\varepsilon)}\right) &\leq \frac{1}{e^{\lambda n(p+\varepsilon)}} \mathbb{E}[e^{\lambda(X_1+\ldots+X_n)}] && && (1) \\
&= \frac{1}{e^{\lambda n(p+\varepsilon)}} e^{h(\lambda)} && \text{per 2.ii} && (2) \\
&= \frac{1}{e^{\lambda n(p+\varepsilon)}} e^{n\mu(\lambda)} && \text{per 2.ii} && (3) \\
&= e^{n\mu(\lambda)-\lambda n(p+\varepsilon)} && && (4) \\
&= e^{n(\mu(\lambda)-\lambda(\varepsilon+p))} && && (5)
\end{aligned}$$

□

3. ..

(i) What are the marginal distributions of X and Y?

$$\begin{aligned}
\mathbb{P}(X=0) &= \sum_{y \in Y} \mathbb{P}(X=0, Y=y) = 0 + 1/4 + 1/4 = 1/2 \\
\mathbb{P}(X=1) &= \sum_{y \in Y} \mathbb{P}(X=1, Y=y) = 1/4 \\
\mathbb{P}(X=-1) &= \sum_{y \in Y} \mathbb{P}(X=-1, Y=y) = 1/4
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(Y=0) &= \sum_{x \in X} \mathbb{P}(X=x, Y=0) = 0 + 1/4 + 1/4 = 1/2 \\
\mathbb{P}(Y=1) &= \sum_{x \in X} \mathbb{P}(X=x, Y=1) = 1/4 \\
\mathbb{P}(Y=-1) &= \sum_{x \in X} \mathbb{P}(X=x, Y=-1) = 1/4
\end{aligned}$$

(ii) Are X and Y independent?

No. As a counterexample of independence, note that $\mathbb{P}(X=0, Y=0) = 0$ but $\mathbb{P}(X=0) * \mathbb{P}(Y=0) = (1/2) * (1/2) = 1/4$.

(iii) Compute the covariance $\text{Cov}(X, Y)$

Generally, $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. In this case, the joint probability of X and Y is only nonzero when $XY = 0$, so $\mathbb{E}[XY] = 0$. In addition,

$$\mathbb{E}[X] = \sum_{x \in R(X)} x\mathbb{P}(X = x) = -1(1/4) + 1(1/4) + 0(1/2) = 0$$

And symmetricaly, $\mathbb{E}[Y] = 0$. Thus, because $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0$, $\boxed{\text{Cov}(X, Y) = 0}$.

4. (i) What is the joint distribution of B and X_l?
By the definition of conditional probability,

$$\mathbb{P}(B = b, X_L = x) = \mathbb{P}(X_L = x | B = b) * \mathbb{P}(B = b) \tag{6}$$

If we are drawing from box b, b/n balls are red. The number of red balls we select in L draws with replacement is thus a binomial random $X_L | B = b \sim \text{Binomial}(m, b/n)$.

$$\mathbb{P}(X_l = x | B = b) = \binom{L}{x} \left(\frac{b}{n}\right)^x \left(1 - \frac{b}{n}\right)^{L-x} \tag{7}$$

Noting that $\mathbb{P}(B = b) = 1/n$ and plugging (7) into (6), we conclude

$$\mathbb{P}(B = b, X_L = x) = \frac{1}{n} \binom{L}{x} \left(\frac{b}{n}\right)^x \left(1 - \frac{b}{n}\right)^{L-x}$$

(ii) What is the conditional distribution of B given $X_L = j$? Per Baye's Rule,

$$\mathbb{P}(B = b | X_L = j) = \frac{\mathbb{P}(X_L = j | B = b)\mathbb{P}(B = b)}{\mathbb{P}(X_L = j)}$$

We worked the numerator out in 4i. The denominator is given by the partition rule:

$$\mathbb{P}(B = b | X_L = j) = \frac{\frac{1}{n}\binom{L}{j}\left(\frac{b}{n}\right)^j \left(1 - \frac{b}{n}\right)^{L-j}}{\sum_{b=1}^{n} \mathbb{P}(X_L = j | B = b) * \mathbb{P}(B = b)}$$

$\mathbb{P}(B = b)$ in the denominator $= 1/n$. Since the same term exists in the numerator, we can cancel both. Further expanding the denominator, we find

$$\mathbb{P}(B = b | X_L = j) = \frac{\binom{L}{j}\left(\frac{b}{n}\right)^j \left(1 - \frac{b}{n}\right)^{L-j}}{\sum_{b=1}^{n} \binom{L}{j}\left(\frac{b}{n}\right)^j \left(1 - \frac{b}{n}\right)^{L-j}}$$

Cancelling the $\binom{L}{x}$ terms, we conclude

$$\mathbb{P}(B = b | X_L = j) = \frac{\left(\frac{b}{n}\right)^j \left(1 - \frac{b}{n}\right)^{L-j}}{\sum_{b=1}^{n} \left(\frac{b}{n}\right)^j \left(1 - \frac{b}{n}\right)^{L-j}}$$

(iii) What value of k maximizes the conditional distribution from part (ii)?

Maximizing the conditional probability in 4ii is equivalent to maximizing its numerator. We can find the value of k that maximizes the numerator by taking the numerator's derivative with respect to k, then setting that expression equal to 0 to find critical points.

$$\begin{aligned}
\frac{d}{dk}\left(\left(\frac{k}{n}\right)^j \left(1 - \frac{k}{n}\right)^{L-j}\right) &= \frac{j}{n}\left(\frac{k}{n}\right)^{j-1}\left(1 - \frac{k}{n}\right)^{L-j} + \left(\frac{k}{n}\right)^j \frac{-(L-j)}{n}\left(1 - \frac{k}{n}\right)^{L-j-1} \\
&= \frac{1}{n}\left(\frac{k}{n}\right)^{j-1}\left(1 - \frac{k}{n}\right)^{L-j-1}\left[j\left(1 - \frac{k}{n}\right) - \frac{k}{n}(L-j)\right] \\
&= \frac{1}{n}\left(\frac{k}{n}\right)^{j-1}\left(1 - \frac{k}{n}\right)^{L-j-1}\left(j - \cancel{\frac{jk}{n}} - \frac{kL}{n} + \cancel{\frac{kj}{n}}\right) \\
&= \frac{1}{n}\left(\frac{k}{n}\right)^{j-1}\left(1 - \frac{k}{n}\right)^{L-j-1}\left(j - \frac{kL}{n}\right)
\end{aligned}$$

Setting the derivative equal to 0, we see that $k = \frac{jn}{L}$ is a critical point, as are $k = n$ and $k = 0$, though neither of the latter two makes sense as a maximum.

$\boxed{\text{We thus conclude that } k = \frac{jn}{L} \text{ maximizes the conditional probability in 4ii.}}$

5. Let S_n be a simple random walk on the integers, starting from $S_0 = 0$. Let M_n denote the maximum of $S_0, S_1, \ldots, S_n$. Let $N_n^{max}(b)$ denote the number of paths for which $S_k \geq b$ at some time $k \in \{1, \ldots, n\}$. Let $N_n^+(b)$ be the number of paths for which $S_n > b$. Assume b is an even integer and n is odd so that $\mathbb{P}(S_n = b) = 0$.

(i) Explain why $N_n^{max}(b) = 2 \cdot N_n^+(b)$.

Any path in $N_n^+(b)$ must have some point $S_k = b$ and $S_{k+1} = b + 1$. Mirroring the reflection principle, imagine an alternate path S^* such that ${S^*}_{k+j} = b - (S_{k+j} - b)$ for $j \in \{k+1 \ldots n\}$. i.e., an alternate path that is the same as the original up to time k where the path hits b, then for time $t \in \{K+1, \ldots, n\}$ is the reflection of the original path over b.

Since every path constituting those counted in $N_n^+(b)$ has such an alternate path, $N_n^{max}(b) = 2 \cdot N_n^+(b)$.

(ii) Write a formula for $\mathbb{P}(M_n \geq b)$ in terms of n and b.

$$\mathbb{P}(M_n \geq b) = \frac{N_n^{max}(b)}{\#\text{ paths}} \tag{8}$$

Per 5i, this is equivalent to

$$\mathbb{P}(M_n \geq b) = \frac{2N_n^+(b)}{\#\text{ paths}} = \frac{2N_n^+(b)}{2^n} = 2^{n-1} N_n^+(b) \tag{9}$$

$N_n^+(b)$ is the number of paths with $S_n \geq b$. Imagine that the direction of each step of our random walk is determined by the toss of a coin. For a walk to have $S_n \geq b$,

$$\begin{aligned} heads - tails &\geq b \\ heads &\geq b + tails \\ heads &\geq b + (n - heads) \\ 2 * heads &\geq b + n \\ heads &\geq \frac{b+n}{2} \end{aligned}$$

The number of ways to flip k heads in n tosses is given by $\binom{n}{k}$. Any number of heads between $(b+n)/2$ and n will give $S_n \geq b$. Thus,

$$N_n^+(b) = \sum_{k=(b+n)/2}^{n} \binom{n}{k} \tag{10}$$

Finally, plugging this into (9), we conclude

$$\mathbb{P}(M_n \geq b) = \frac{\sum_{k=(b+n)/2}^{n} \binom{n}{k}}{2^{n-1}}$$