Homework 2: Due 6/5/14

1. Let X be a random variable with binomial distribution $b(n, p)$, which means that the space of X is $S = \{0, 1, 2, \dots, n\}$ and the p.m.f. $f(x)$ is $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$. Show that $E(X) = np$ and $E\left((X-\mu)^2\right) = np(1-p)$. You may use without proof that $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ (the *Binomial Theorem*), and I would suggest you show that $E\left((X-\mu)^2\right) = E(X(X-1)) + E(X) - \mu^2$ and examine the expression $E(X(X-1))$.

Solution: We are asked to show that $E(X) = np$, where the p.m.f. of X is $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ and the space of X is $S = \{0, 1, 2, \dots, n\}$. We compute $E(X)$ as follows, using $q := 1-p$.

$$\begin{aligned}
E(X) &= \sum_{x=0}^{n} x f(x) = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} && \because \textit{ zero-th term is zero} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x} = \sum_{k=0}^{n-1} \frac{n!}{k!(n-k-1)!} p^{k+1} q^{n-k-1} && \because \textit{ set } k = x-1 \\
&= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} p^k q^{n-1-k} = np(p+q)^{n-1} && \because \textit{ Binomial Theorem} \\
&= np && \because p+q = 1.
\end{aligned}$$

Next we are asked to show that $\sigma^2 = npq$. We begin by showing that the identity $E\left(X^2 - \mu^2\right) = E\left(X(X-1)\right) + E(X) - \mu^2$ holds.

$$E\left(X(X-1)\right) + E(X) - \mu^2 = E(X^2 - X) + E(X) - \mu^2 = E(X^2 - X + X - \mu^2) = E(X^2 - \mu^2).$$

Next, we examine the term $E\left(X(X-1)\right)$.

$$\begin{aligned}
E\left(X(X-1)\right) &= \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x} = \sum_{x=2}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x} && \because \textit{ first two terms zero} \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x q^{n-x} = \sum_{k=0}^{n-2} \frac{n!}{k!(n-k-2)!} p^{k+2} q^{n-k-2} && \because \textit{ set } k = x-2 \\
&= n(n-1)(n-2)p^2 \sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-2-k)!} p^k q^{n-2-k} \\
&= n(n-1)(n-2)p^2 (p+q)^{n-2} && \because \textit{ Binomial Theorem} \\
&= n(n-1)p^2 && \because p+q = 1.
\end{aligned}$$

So, via the identity we showed above and the fact we computed that $E(X) = np$, we have $\sigma^2 = E\left(X(X-1)\right) + E(X) - \mu^2 = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p) = npq$, as desired.

$\square$

2. On a 8 question multiple choice exam there are five possible answers (a), (b), (c), (d), and (e) for each question, and exactly one answer is correct. What is the probability mass function of X the number of correct answers if a student answers randomly? What is the probability that the student who answers randomly gets the first and last question right, and every other question wrong? What is the probability that they get exactly 3 questions correct?

Solution: This is a binomial distribution with $n = 8$ and $p = .2$, thus the p.m.f. $f(x)$ is

$$f(x) = \binom{8}{x}(.2)^x(.8)^{8-x}.$$

The probability they get the first and last question wrong is $(.2)(.8)\cdots(.8)(.2) = (.2)^2(.8)^6 =$, and corresponds to 8 Bernoulli trials with $p = .2$ and 2 successes. The probability that the student gets exactly 3 questions correct is $f(3) = \binom{8}{3}(.2)^3(.8)^5$.

□

3. Using X from the Problem 2 what is the moment generating function of X? Show that the m.g.f. $M(t)$ satisfies $M'(0) = E(X)$ and $M''(0) = E(X^2)$. Find μ, σ, and σ^2 for this random variable using the formula from Problem 1.

Solution: The m.g.f. of X is

$$M(t) = \sum_{x=1}^{8} e^{tx} f(x) = \sum_{x=1}^{8} e^{tx}\binom{8}{x}(.2)^x(.8)^{8-x}.$$

We have that $M'(0) = \mu$, and $M''(0) = \sigma^2$. We calculate $M'(t)$ and $M''(t)$ first.

$$M'(t) = \sum_{x=1}^{8}(e^{tx})'f(x) = \sum_{x=1}^{8} xe^{tx}f(x), \text{ and}$$
$$M''(t) = \sum_{x=1}^{8}(e^{tx})''f(x) = \sum_{x=1}^{8} x^2e^{tx}f(x).$$

So substituting $t = 0$ into the above formulas we immediately see that $M'(0) = \mu$ and $M''(0) = \sigma^2$. Now since X is has binomial distribution $b(8, .2)$ we have that $\mu = np = 8(.2) = 1.6$ and $\sigma^2 = npq = 8(.2)(.8) = 1.28$. And $\sigma = \sqrt{1.28} =$.

□

4. Let X be the number of randomly selected people you must ask in order to find someone that has the same birthday as you. Assume each day is equally likely and ignore leap years. What is the p.m.f. of X? What is μ, σ^2 and σ of X? What is the probability that you have to ask more than 400 people? What about fewer than 300?

Solution: This is a geometric distribution. In particular, if it takes x people until you find someone with your birthday, then you failed $x-1$ times and succeeded once in x Bernoulli trials with chance of success $p = 1/365$, hence $f(x) = \left(\frac{364}{365}\right)^{x-1} \cdot \frac{1}{365}$. For a geometric distribution with chance of success p and $q = 1 - p$, we have

$$\begin{aligned} M(t) &= \sum_{x=1}^{\infty} e^{tx} p q^{x-1} \\ &= pe^t \sum_{x=1}^{\infty} (e^t q)^{x-1} \\ &= pe^t \frac{1}{1 - e^t q} = \frac{pe^t}{1 - qe^t}. \end{aligned}$$

So when $p = 1/365$ we have $M(t) = \frac{e^t}{365 - 364e^t}$. Now $\mu = M'(0)$ so we calculate

$$\begin{aligned} M'(t) &= \frac{(365 - 364e^t)e^t - e^t(-364e^t)}{(365 - 364e^t)^2} \\ &= \frac{365e^t - 364e^{2t} + 364e^{2t}}{(365 - 364e^t)^2} = \frac{365e^t}{(365 - 364e^t)^2}, \end{aligned}$$

and plugging in $t = 0$ we get $\mu = M'(0) = 365$. Note that in general $M'(0) = p^{-1}$ for general p in a geometric distribution, which is seen by differentiating $M(t) = \frac{pe^t}{1-qe^t}$ and setting $t = 0$. Finding σ^2 is a little messier, but we use the formula for $M'(t)$ we found in the equation above, to get

$$\begin{aligned} M''(t) = \left(\frac{365e^t}{(365 - 364e^t)^2}\right)' &= \frac{(365 - 364e^t)^2(365e^t) - (365e^t)(2)(\text{-}364e^t)(365 - 364e^t)}{(365 - 364e^t)^4} \\ &= \frac{(365 - 364e^t)365e^t + 2 * 365e^t * 364e^t}{(365 - 364e^t)^3} = \frac{365^2 e^t + 365 * 364e^{2t}}{(365 - 364e^t)^3} \end{aligned}$$

$$M''(0) = \frac{365^2 + 365 * 364}{1} = 365(365 + 364).$$

Hence, $\sigma^2 = 365(365 + 364) - 365^2 = 365 * 364 = 132860$ and $\sigma = \sqrt{365 * 364} = 2\sqrt{33215}$.

Note that for a general geometric distribution we have $\sigma^2 = q/p^2$ which can be seen by differentiating $M(t) = \frac{pe^t}{1-qe^t}$ twice and then setting $t = 0$. Indeed, $E(X^2) = M''(0) = (1 + q)/p^2$, so $\sigma^2 = (1 + q)/p^2 - (1/p)^2 = q/p^2$.

Moving on, the probability that you have to ask more than 400 people is $P(X > 400)$, which is a geometric series.

$$P(X > 400) = \sum_{x=401}^{\infty} f(x) = \sum_{x=401}^{\infty} \frac{1}{365}\left(\frac{364}{365}^{x-1}\right) = \sum_{x=0}^{\infty} \frac{1}{365}\left(\frac{364}{365}\right)^{400}\left(\frac{364}{365}\right)^{x},$$

for which we use the formula $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$, obtaining

$$P(X > 400) = \frac{1}{365}\left(\frac{364}{365}\right)^{400} \frac{1}{1-\frac{364}{365}} = \left(\frac{364}{365}\right)^{400} \frac{1}{365-364} = \left(\frac{364}{365}\right)^{400} \approx .3337.$$

In fact, a similar argument shows that $P(X > x) = \left(\frac{364}{365}\right)^{x}$, so $P(X < 300) = 1 - P(X > 299) = \left(\frac{364}{365}\right)^{299} \approx .5597$.

□

5. Consider X from the problem above except that X is the number of randomly selected people you must ask in order to find n people that have the same birthday as you. Find the p.m.f. $f(x)$ of X and the m.g.f. $M(t)$ of X.

Solution: The p.m.f. $f(x)$ of X is

$$f(x) = \binom{x-1}{n-1}\left(\frac{1}{365}\right)^{n}\left(\frac{364}{365}\right)^{x-n},$$

which is seen by considering that if you have to find n people with the same birthday, then the first time you have succeeded you had previously found $n-1$ people with the same birthday as you in $x-1$ attempts (hence the binomial coefficient, taking into account the various possible locations of the successful finds) and in any particular case you have, at the moment you succeeded in finding the nth person, made n successes and $x-n$ failures in a x attempts of a Bernoulli trial. Now the m.g.f. of X is easily seen to be

$$M(t) = \sum_{x=n}^{\infty} e^{nt}\binom{x-1}{n-1}\left(\frac{1}{365}\right)^{n}\left(\frac{364}{365}\right)^{x-n}.$$

□