2. (Fish)

(a) Solution: This is a problem similar to the Coupon Collector's Problem covered in lecture. Let X_i denote the number of fish caught while we already have $i-1$ different fish caught. Then the total number of fish caught in order to get all kinds is $X = \sum_{i=1}^{n} X_i$ where X_i follows a geometric distribution, i.e, $X_i \sim Geom(p_i)$, and $p_i = 1 - \frac{i-1}{n}$.

$$
\begin{aligned}
\mathbb{E}(X_i) &= \frac{1}{p_i} = \frac{n}{n-i+1} \\
\mathbb{E}(X) &= \mathbb{E}\Big[\sum_{i=1}^{n} X_i\Big] \\
&= \sum_{i=1}^{n} \mathbb{E}[X_i] \\
&= \sum_{i=1}^{n} \frac{n}{n-i+1} \\
&= n\sum_{i=1}^{n} \frac{1}{i} \\
&= nH(n) \approx n(\ln n + \Theta(1)) = n\ln n + \Theta(n)
\end{aligned}
$$

(b) Solution: Let X_i be an indicator random variable denoting whether the fish kind i is not caught when $2n$ fish have been caught. Let X denote the number of kinds of fish not caught when $2n$ fish have been caught.

$$
\begin{aligned}
\mathbb{E}(X_i) &= Pr[X_i = 1] \\
&= \big(1 - \frac{1}{n}\big)^{2n} \\
\mathbb{E}(X) &= \mathbb{E}(\sum_{i=1}^{n} X_i) \\
&= \sum_{i=1}^{n} \mathbb{E}(X_i) \\
&= n\big(1 - \frac{1}{n}\big)^{2n}
\end{aligned}
$$

(c) Solution: Let Y_i be an indicator random variable denoting whether the fish kind i has only been caught exactly once after $3n$ fish being caught.

$$
\begin{aligned}
\mathbb{E}(Y_i) &= Pr[Y_i = 1] \\
&= \binom{3n}{1}(1-\frac{1}{n})^{3n-1}(\frac{1}{n}) \\
\mathbb{E}(Y) &= \mathbb{E}(\sum_i^n Y_i) \\
&= \sum_i^n \mathbb{E}(Y_i) \\
&= n\binom{3n}{1}(1-\frac{1}{n})^{3n-1}(\frac{1}{n}) = 3n(\frac{n-1}{n})^{3n-1}
\end{aligned}
$$

(d) Solution: We can use similar method in (a). Let X_i denote the number of fish caught while we already have $i-1$ different fish caught. Then the total number of fish we need to get $n/2$ kinds of fish is $X = \sum_{i=1}^{n/2} X_i$ where $X_i \sim Geom(1-\frac{i-1}{n})$.

$$
\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(\sum_i^{n/2} X_i) \\
&= \sum_{i=1}^{n/2} \frac{n}{n-i+1} \\
&= n \sum_{i=n/2+1}^{n} \frac{1}{i} \\
&= n(H(n) - H(n/2)) \approx n(\ln n - \ln n/2) + \Theta(n) = n\ln 2 + \Theta(n)
\end{aligned}
$$

(e) Solution: Let F_i be the total number of fish for generation i and F be the number of fish in all generations.

$$
\begin{aligned}
\mathbb{E}(F_1) &= 1p_1 + 2p_2 \\
\mathbb{E}(F_2) &= \mathbb{E}(F_1)p_1 + 2\mathbb{E}(F_1)p_2 = \mathbb{E}(F_1)(p_1+2p_2) = (p_1+2p_2)^2 \\
\mathbb{E}(F_3) &= \mathbb{E}(F_2)p_1 + 2\mathbb{E}(F_2)p_2 = \mathbb{E}(F_2)(p_1+2p_2) = (p_1+2p_2)^3 \\
&\vdots \\
\mathbb{E}(F_i) &= (p_1+2p_2)^i
\end{aligned}
$$

Thus, the expected number of fish in the talk will be

$$
\begin{aligned}
\mathbb{E}(F) &= \mathbb{E}(\sum_{i=1}^{\infty} F_i) \\
&= \sum_{i=1}^{\infty}(p_1+2p_2)^i = \frac{1}{1-(p_1+2p_2)}, \text{ given } p_1+2p_2<1
\end{aligned}
$$

The above sum is bounded only when $p_1 + 2p_2 < 1$.