# SPECIAL DISCRETE DISTRIBUTIONS

**Bernoulli**

Consider an experiment with only two possible outcomes (`Success` and `Failure`) which happen with the probability of $p$ and $q \equiv 1-p$ respectively

We define a random variable $X$ as the number of successes one gets. Its distribution is obviously

| $X =$ | 0 | 1 |
|---|---|---|
| Prob: | $q$ | $p$ |

implying: $\mathbb{E}(X) = p$, $\mathrm{Var}(X) = p - p^2 = pq$, $P(z) = q + pz$.

**Binomial**

The experiment consists of $n$ *independent* rounds (trials) of the Bernoulli type. The sample space consists of all $n$-letter words built of letters $S$ and $F$. These are *not* equally likely, the probability of each is $p^i q^{n-i}$ where $i$ is the number of $S$'s and $n-i$ is the number of $F$'s. We also know that there is $\binom{n}{i}$ 'words' with exactly $i$ $S$'s. Thus, the probability that $X$, the *total number of successes*, will have the value of $i$ is

$$\binom{n}{i} p^i q^{n-i}$$

where $i = 0,\ 1,\ 2, \ldots\ , n-1,\ n$. They do add up to 1: $\sum\limits_{i=0}^{n} \binom{n}{i} p^i q^{n-i} = (p+q)^n$ (check). Symbolically, we denote this distribution $\mathcal{B}(n,p)$. Clearly, it has two *parameters*, $n$ and $p$. There are three ways of deriving $\mathbb{E}(X)$:

1. 'Algebraically', by struggling with $\sum\limits_{i=0}^{n} i \cdot \binom{n}{i} p^i q^{n-i}$.

2. 'Statistical' argument: $X \equiv X_1 + X_2 + \ldots.. + X_n$, where $X_i$ are independent, identically distributed (IID), all having the Bernoulli distribution. Thus: $\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \ldots.. + \mathbb{E}(X_n) = p + p + \ldots. + p = \boxed{np}$

3. Using PGF, which, for the same reason equals $(q+pz)\cdot(q+pz)\cdot\ldots\cdot(q+pz) = (q+pz)^n$.

   One differentiation yields: $n(q+pz)^{n-1}p\big|_{z=1} = np$.

Similarly, we can compute Var($X$), by using the 'statistical' approach (easiest): Var($X$) =Var($X_1$)+Var(
...+Var($X_n$) $= pq + pq + \ldots + pq = \boxed{npq}$

There is no 'compact' formula for the distribution function $F(i)$. This means that the probability of a *range* of values can be computed only by adding the individual probabilities. For example, if $n = 20$ and $p = \frac{1}{6}$, $\Pr(X \geq 10) = \binom{20}{10}(\frac{1}{6})^{10}(\frac{5}{6})^{10} + \binom{20}{11}(\frac{1}{6})^{11}(\frac{5}{6})^{9} + \binom{20}{12}(\frac{1}{6})^{12}(\frac{5}{6})^{8} + \ldots + \binom{20}{0}(\frac{1}{6})^{20}(\frac{5}{6})^{0} = 0.05985\%$.

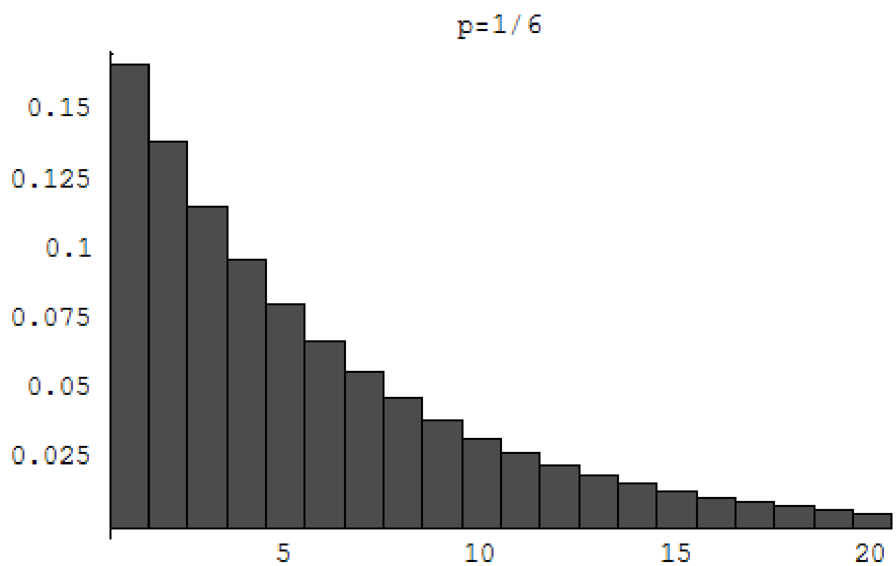
Your main task will be first to *recognize* a binomial RV when you see one.

**Geometric**

This distribution is based on the same kind of experiment, where the trials are performed, repeatedly, until the *first success* appears. The random variable (let's call it $X$ again) is now the total *number of trials*. The sample space looks like this: $S$, $FS$, $FFS$, $FFFS, ...$ (infinite), the individual probabilities are: $p$, $qp$, $q^2p$, $q^3p, ...$ In general, we have

$$\Pr(X = i) = pq^{i-1}$$

where $i = 1,\ 2,\ 3, ...$ Again, they do add up to 1: $p(1 + q + q^2 + q^3 + ...) = \frac{p}{1-q} = \frac{p}{p} = 1$. Symbolic notation: $\mathcal{G}(p)$, having *one* parameter $p$. Its histogram looks like this:




To compute $\mathbb{E}(X)$, we can try the algebraic approach: $p(1 + 2q + 3q^2 + 4q^3 + ...)$, which is not so easy. On the other hand, the PGF equals: $\sum_{i=1}^{\infty} z^i \cdot pq^{i-1} = zp \sum_{i=1}^{\infty} (zq)^{i-1} = pz[1 + zq + (zq)^2 + (zq)^3 + ...] = \dfrac{pz}{1-qz}$. Differentiating with respect to $z$: $\frac{p}{(1-qz)^2}\Big|_{z=1} = \frac{p}{p^2} = \boxed{\dfrac{1}{p}}$. One more differentiation yields: $\frac{2pq}{(1-qz)^3}\Big|_{z=1} = \frac{2q}{p^2}$, implying:

$$\mathrm{Var}(X) = \tfrac{2(1-p)}{p^2} + \frac{1}{p} - (\tfrac{1}{p})^2 = \tfrac{(1-p)}{p^2} = \tfrac{1}{p}\left(\tfrac{1}{p} - 1\right)$$

The standard deviation is the square root of this,.e.g. the number of trials to get the first 6 has the mean of 6 and standard deviation equal to $\sqrt{6 \times 5} = 5.477$.

This time it is easy to find the distribution function $F(i)$. We first compute $\Pr(X > i) = pq^i + pq^{i+1} + pq^{i+2} + ... = pq^i(1 + q + q^2 + ....) = \frac{pq^i}{1-q} = q^i$, resulting in

$$F(i) = \Pr(X \leq i) = 1 - q^i$$

where $i = 1, 2, 3, ....$ Thus, for example, the probability that it will take *at least* 10 rolls to get the first 6 is $1 - F(9) = (\frac{5}{6})^9 = 19.38\%$. Similarly, the probability that it will take more than 18 rolls is $(\frac{5}{6})^{18} = 3.756\%$.

The geometric distribution is `memoryless`: Given that $X > k$, the conditional distribution of $X - k$ is the same $\mathcal{G}(p)$ of the original $X$.

**Negative Binomial**

distribution is a simple extension of the *geometric* distribution, where $X$ is now the *number of trials* till (and including) the $k^{th}$ success. Obviously, it can be expressed a sum of $k$ *independent* random variables of the *geometric* type: $X = X_1 + X_2 + .... + X_k$. This yields immediately (the 'statistical' argument)

$$\mathbb{E}(\mathbb{X}) = \frac{k}{p}$$

and

$$\text{Var}(X) = \frac{k}{p} \cdot (\frac{1}{p} - 1)$$

Similarly, the new PGF is obtained by raising the old (geometric) PGF to the power of $k$: $(\frac{pz}{1-qz})^k$. The symbolic name will be $\mathcal{NB}(k, p)$, there are clearly two parameters, $k$ and $p$.

To get the individual probabilities $\Pr(X = i)$ takes a bit more ingenuity. To get the $k^{th}$ success in the $i^{th}$ roll, we need: (i) $k - 1$ successes, in *any order*, during the first $i - 1$ rolls, and (ii) a success in the $k^{th}$ roll.. The probability of this happening is: $\binom{i-1}{k-1} p^{k-1} q^{i-k} \cdot p =$

$$\binom{i-1}{k-1} p^k q^{i-k} \equiv \binom{i-1}{i-k} p^k q^{i-k}$$

where $i = k,\ k + 1,\ k + 2,$ ... It helps to display these in an explicit table:

| $X =$ | $k$ | $k+1$ | $k+2$ | $k+3$ | ... |
|---|---|---|---|---|---|
| Prob: | $p^k$ | $kp^kq$ | $\binom{k+1}{2}p^kq^2$ | $\binom{k+2}{3}p^kq^3$ | ... |

To verify that they add up to 1, we proceed as follows: $1 \equiv p^k(1-q)^{-k} = p^k \left[1 - \binom{-k}{1}q + \binom{-k}{2}q^2 - \binom{-k}{3}q^3 + \ldots\right.$ $p^k \left[1 + kq + \binom{k+1}{2}q^2 + \binom{k+2}{3}q^3 + ...\right]$.

We can now easily deal with questions like: what is the probability that it will take exactly 5 flips of a coin to get the third head: $\binom{4}{2}(\frac{1}{2})^3(\frac{1}{2})^{5-3} = 18.75\%$, and: what is the probability of requiring exactly 10 rolls to get the second six: $\binom{9}{1}(\frac{1}{6})^2(\frac{5}{6})^8 = 5.814\%$].

To be able to answer questions like: 'what is the probability that we will need *more than* 10 rolls to get a second 6', we realize that this is the same as the probability of getting, in the first 10 rolls, *fewer* than 2 (i.e. 0 or 1) sixes, i.e.: $(\frac{5}{6})^{10} + \binom{10}{1}(\frac{5}{6})^9(\frac{1}{6})^1 = 48.45\%$.

In general,

$$\Pr(X > i) = \sum_{j=0}^{k-1} \binom{i}{j} p^j q^{i-j}$$

**Hypergeometric**

distribution relates to the following experiment: Suppose there are $N$ objects, $K$ of which have some *special* property (such as being red marbles, spades, aces, defective items, women), etc., the remaining $N-K$ objects will be called '*ordinary*'. Of these $N$ objects, $n$ are *randomly* selected, without replacement. Let $X$ be the number of 'special' objects found in the sample.

The sample space consists of a list of all possible ways of selecting $n$ objects out of $N$ (order irrelevant). We know that the *total* number of these is $\binom{N}{n}$ and that they are equally likely. We also know (having solved many questions of this type) that $\binom{K}{i} \cdot \binom{N-K}{n-i}$ of these 'simple events' contains exactly $i$ 'special' objects. Thus

$$\Pr(X = i) = \frac{\binom{K}{i} \cdot \binom{N-K}{n-i}}{\binom{N}{n}}$$

where $\max(0, n-N+K) \le i \le \min(n, K)$. The formula which verifies that these probabilities add up to 1 is called hypergeometric (we must skip this).

Note that when the sampling is done *with replacement*, the correct distribution is binomial (with $p = \frac{K}{N}$).

To derive $\mathbb{E}(X)$, we use the 'statistical' approach: $X = X_1 + X_2 + ... + X_n$, where $X_i$ is the number of 'special' objects obtained in the $i^{th}$ 'draw' (clearly, of Bernoulli type, with $p = \frac{K}{N}$). Note that this time the $X_i$'s are *not* independent!

The expected value of $X$ is easy: $\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \ldots.. + \mathbb{E}(X_n) = \frac{K}{N} + \frac{K}{N} + \ldots + \frac{K}{N} =$

$$n \cdot \frac{K}{N}$$

(an exact analog of the binomial $np$ formula).

To establish $\mathrm{Var}(X)$, we use: $\sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i, X_j) = n \cdot \mathrm{Var}(X_1) + 2\binom{n}{2} \cdot \mathrm{Cov}(X_1, X_2)$ as all variances, and all covariances, must have the same value. We already know that $\mathrm{Var}(X_i) = \frac{K}{N} \cdot \frac{N-K}{N}$ (the $pq$ formula of Bernoulli distribution). To find $\mathrm{Cov}(X_1, X_2)$ we first build their joint distribution:

| $X_1=$ / $X_2=$ | 0 | 1 |
|---|---|---|
| 0 | $\frac{N-K}{N} \cdot \frac{N-K-1}{N-1}$ | $\frac{K}{N} \cdot \frac{N-K}{N-1}$ |
| 1 | $\frac{N-K}{N} \cdot \frac{K}{N-1}$ | $\frac{K}{N} \cdot \frac{K-1}{N-1}$ |

from which it easily follows that $\mathbb{E}(X_1 \cdot X_2) = \frac{K\,(K-1)}{N\,(N-1)}$. This implies: $\mathrm{Cov}(X_1, X_2) = \frac{K\,(K-1)}{N\,(N-1)} - \left(\frac{K}{N}\right)^2 = -\frac{K\,(N-K)}{N^2(N-1)}$. This enables us to complete the previous computation:

$$\begin{aligned} \mathrm{Var}(X) &= n\frac{K\,(N-K)}{N^2} - n(n-1)\frac{K\,(N-K)}{N^2(N-1)} \\ &= n \cdot \frac{K}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1} \end{aligned}$$

The first three factors are the analog of $npq$, reduced by a 'correction factor' of $\frac{N-n}{N-1}$. Note that, when $n = 1$, this factor is equal to 1 (check), and when $n = N$, it becomes 0 ($X$ has then a `degenerate distribution` with only one possible value).

Symbolically, we denote this distribution $\mathcal{HG}(N, K, n)$.

EXAMPLE: There are 30 red and 70 blue marbles in a box. If 10 marbles are randomly drawn (without replacement), what is the probability that exactly 4 of these are red? Answer: $\frac{\binom{30}{4}\times\binom{70}{6}}{\binom{100}{10}} = 20.76\%$. *With replacement*: $\binom{10}{4} \times 0.3^4 \times 0.7^6 = 20.01\%$.

**Poisson**

distribution relates to the following type of experiment: Suppose customers arrive (to a gas station), randomly, at a steady rate of $\lambda$ (say 1.2) per minute. Let $X$ be the *number of customers* who will arrive during the next $T$ (say 3) minutes. Based on this, the mean value of $X$ should be $\lambda T \equiv \Lambda$ (=3.6).

We start by dividing $T$ into $n$ subintervals (36 intervals of 5 seconds, say), taking the probability of a *single* arrival during each subinterval to be $p = \frac{\Lambda}{n}$, independent of what happens in other subintervals. With this kind of model, $X$ has the binomial distribution with $n$ and $p$, implying that

$$\Pr(X = i) = \binom{n}{i}\left(\frac{\Lambda}{n}\right)^i\left(1 - \frac{\Lambda}{n}\right)^{n-i}$$

where $i = 0, 1, 2, ..., n$ (note that $X$ has the correct mean). What's wrong with this model, and how do we fix it? Well, in a real situation, more than one customer can arrive during 5 seconds. To prevent this from happening, we increase the value of $n$. Only in the $n \to \infty$ limit, we reach the perfect answer: $\lim_{n\to\infty} \binom{n}{i}\left(\frac{\Lambda}{n}\right)^i\left(1 - \frac{\Lambda}{n}\right)^{n-i} = \frac{\Lambda^i}{i!}\lim_{k\to\infty} \frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdot\ldots\cdot\frac{n-i+1}{n}\cdot\left(1 - \frac{\Lambda}{n}\right)^{n-i} = \frac{\Lambda^i}{i!}\cdot\lim_{k\to\infty}\left(1 - \frac{\Lambda}{n}\right)^{n-i} = \frac{\Lambda^i}{i!}\cdot e^{-\Lambda}$. Thus

$$\Pr(X = i) = \frac{\Lambda^i}{i!}e^{-\Lambda}$$

where $i = 0, 1, 2,$ ... (all non-negative integers). We can easily verify that these probabilities add up to 1: $\left(1 + \Lambda + \frac{\Lambda^2}{2} + \frac{\Lambda^3}{3!} + ..\right)e^{-\Lambda} = e^{\Lambda}\cdot e^{-\Lambda} = 1$. The corresponding PGF is $e^{-\Lambda}\sum_{i=0}^{\infty}\frac{\Lambda^i}{i!}z^i = e^{\Lambda(z-1)}$. Differentiating the PGF with respect to $z$: $\Lambda\exp[\Lambda(z-1)]|_{z=1} = \Lambda$, as expected. One more differentiation yields: $\Lambda^2 \exp[\Lambda(z-1)]|_{z=1} = \Lambda^2$. This implies that $\mathrm{Var}(X) = \Lambda^2 + \Lambda - \Lambda^2 = \Lambda$ (same as the mean!).

$\mathcal{PS}(\Lambda)$ will be our symbolic notation for this distribution, which has only *one* parameter!

If we have two independent RVs, both of the Poisson type, say $X_1 \in \mathcal{PS}(\Lambda_1)$ and $X_2 \in \mathcal{PS}(\Lambda_2)$, the PGF of $X_1 + X_2$ is clearly $e^{(\Lambda_1+\Lambda_2)(z-1)}$. Can you identify the resulting distribution?

EXAMPLE: Customers arrive at an *average* rate of 3.7/hour.

- What is the probability of exactly 1 arrival during the next 15 min.? $\Lambda = T\lambda = \frac{1}{4} \times 3.7 = 0.925$ (make sure to use *same units*). Answer: $e^{-0.925} \times \frac{0.925}{1!} = 36.68\%$.
- What is the probability of *at least* 4 arrival during the next *30 min.* Answer: $1 - \Pr(X \le 3) = 1 - (1 + 1.85 + \frac{1.85^2}{2} + \frac{1.85^3}{6})e^{-1.85} = 11.69\%$.
- If the store opens at 8:00 what is the probability that their *second* customer arrives between 8:20 and 8:45? Define $A$: at least two arrivals by 8:45, and $B$: at least two arrivals by 8:20. We need $\Pr(A \cap \bar{B}) = \Pr(A) - \Pr(B)$ since $B \subset A$. Answer: $\left[1 - e^{-\Lambda_1}(1+\Lambda_1)\right] - \left[1 - e^{-\Lambda_2}(1+\Lambda_2)\right]$ where $\Lambda_1 = \frac{3}{4}\times 3.7 = 2.775$ and $\Lambda_2 = \frac{1}{3}\times 3.7 = 1.2\bar{3}$, i.e. $e^{-1.2\bar{3}} \times 2.2\bar{3} - e^{-2.775} \times 3.775 = 41.52\%$.

## Multivariate distributions

### Multinomial

is an extension of the binomial distribution, in which each trial can result in 3 (or more) possible outcomes (say $W$, $L$, $T$). The trials are still repeated, *independently*, $n$ times, but now we need *three* RVs $X$, $Y$ and $Z$, which count the total number of outcomes each type, respectively.

The sample space consists of $3^n$ simple events, there are exactly $\binom{n}{i,j,k}$ of these having exactly $i$ $W$'s, $j$ $L$'s and $k$ $T$'s, each of them having the same probability of $p_x^i\, p_y^j\, p_z^k$, implying

$$\Pr(X = i \cap Y = j \cap Z = k) = \binom{n}{i,j,k} p_x^i\, p_y^j\, p_z^k$$

for any non-negative integers $i$, $j$, $k$ which add up to $n$. This formula can be easily extended to the case of 4 or more possible outcomes.

Each *marginal* distribution is obviously *binomial*.

EXAMPLES:

- A team plays a series of 10 games. The probability of winning a game is 0.40, losing a game: 0.55, and tying a game: 0.05. What is the probability of finishing with 5 wins, 4 losses and 1 tie? Answer: $\binom{10}{5,4,1} \times 0.4^5 \times 0.55^4 \times 0.05 = 5.90\%$. Supplementary: What is the probability that they win the series (more wins than losses)? Answer: $\Pr(X > 5) + \Pr(X = 5) - \Pr(X = 5 \cap Y = 5) + \Pr(X = 4) - \Pr(X = 4 \cap Y \geq 4) + \Pr(X = 3 \cap Y < 3) + \Pr(X = 2 \cap Y < 2)\ \Pr(X = 1 \cap Y = 0) = 23.94\%$.

- Roll a die 18 times, what is the probability of getting 3 ones, twos, ..., sixes (the most likely outcome)? Answer: $\binom{18}{3,3,3,3,3,3}(\frac{1}{6})^{18} = 0.135\%$. Why is it so small?

Now, what is the *covariance* between $X$ and $Y$. We use our 'statistical' approach: $X = X_1 + X_2 + \ldots. + X_n$ and $Y = Y_1 + Y_2 + \ldots. + Y_n$, so that $\mathrm{Cov}(X, Y) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j) = \sum_{i=1}^{n} \mathrm{Cov}(X_i, Y_i) + \sum_{i \neq j}^{n} \mathrm{Cov}(X_i, Y_j)$. The second set of covariances yields zero (independence), to get the first, we need $\mathrm{Cov}(X_1, Y_1)$, since the rest have the same value. Let's build a table of the distribution:

| $Y_1=$ \ $X_1=$ | 0 | 1 |
|---|---|---|
| 0 | $p_z$ | $p_x$ |
| 1 | $p_y$ | 0 |

which implies: $\mathrm{Cov}(X_1, Y_1) = \mathbb{E}(X_1 \cdot Y_1) - \mathbb{E}(X_1) \cdot \mathbb{E}(Y_1) =$

$0 - p_x \cdot p_y = -p_x p_y$. Thus, finally:

$$\text{Cov}(X, Y) = -np_x\, p_y$$

EXAMPLES:

- Continuing the first example, what is the covariance between the number of wins and the number of losses (in a 10 game series)? Answer: $-10 \times 0.40 \times 0.55 = -2.2$ .

- Rolling a die 18 times, what is the covariance between the number of 3's and the number of 6's obtained? Answer: $-18 \times \frac{1}{6} \times \frac{1}{6} = -0.5$ .

- 10 dice are rolled and we are paid \$5 for each six, but have to pay \$6 for each one. What is the expected value and the standard deviation of our net win? Introduce $X$ for the (total) number of 6's and $Y$ for the number of 1's, our net win is $5X - 6Y$. Its expected value is $5 \times 10 \times \frac{1}{6} - 6 \times 10 \times \frac{1}{6} = -1.\bar{6}$, its variance equals: $5^2 \times 10 \times \frac{1}{6} \times \frac{5}{6} + (-6)^2 \times 10 \times \frac{1}{6} \times \frac{5}{6} + 2 \times 5 \times (-6) \times (-10) \times \frac{1}{6} \times \frac{1}{6} = 101.3\bar{8}$. Answer: $-1.667 \pm 10.069$ dollars.

- A die is rolled 18 times, $U$ is the number of 'small' outcomes (meaning $\leq 3$), $V$ is the number of even outcomes (2, 4 and 6). Find $\text{Cov}(U, V)$. Introduce $U_0 = U - T$ and $V_0 = V - T$, where $T$ counts the 'overlap' outcomes (number of 2's, in this case). Then $\text{Cov}(U, V) = \text{Cov}(U_0 + T, V_0 + T) = \text{Cov}(U_0, V_0) + \text{Cov}(U_0, T) + \text{Cov}(T, V_0) + \text{Var}(T) = -n(p_u - p_t)(p_v - p_t) - n(p_u - p_t)p_t - np_t(p_v - p_t) + np_t(1 - p_t) = -n(p_u p_v - p_t) \equiv -n(p_u p_v - p_{uv})$. Answer: $-18 \times (\frac{1}{2} \times \frac{1}{2} - \frac{1}{6}) = -1.5$

**Multivariate Hypergeometric**

is an extension of the hypergeometric distribution, when having thee (or more) types of objects, e.g. red, blue and green marbles or hearts, diamonds, spades and clubs, etc. We now assume that the total number of objects of each type is $K_1$, $K_2$ and $K_3$ where $K_1+K_2+K_3 = N$. The sample space will still consist of $\binom{N}{n}$ possible selections of $n$ of these, all equally likely. We also know that $\binom{K_1}{i} \times \binom{K_2}{j} \times \binom{K_3}{k}$ of these will contain exactly $i$ objects of Type 1, $j$ objects of Type 2 and $k$ objects of Type 3. Thus,

$$\Pr(X = i \cap Y = j \cap Z = k) = \frac{\binom{K_1}{i}\binom{K_2}{j}\binom{K_3}{k}}{\binom{N}{n}}$$

where $i$, $j$, $k$ are non-negative integers not bigger than $K_1$, $K_2$ and $K_3$ respectively, which add up to $n$.

The *marginal* distribution of $X$ ($Y$, $Z$) is the ordinary hypergeometric with obvious parameters.

The **covariance** between $X$ and $Y$ again follows by taking $X = X_1 + X_2 + ... + X_n$ and $Y = Y_1 + Y_2 + ... + Y_n$: $\text{Cov}(X,Y) = \sum_{i=1}^{n}\text{Cov}(X_i, Y_i) + \sum_{i \neq j}^{n}\text{Cov}(X_i, Y_j) = n \times \text{Cov}(X_1, Y_1) + n(n-1) \times \text{Cov}(X_1, Y_2)$. Now we need the joint distribution of $X_1$ and $Y_1$:

| $X_1=$ / $Y_1=$ | 0 | 1 |
|---|---|---|
| 0 | $\frac{N-K_1-K_2}{N}$ | $\frac{K_1}{N}$ |
| 1 | $\frac{K_2}{N}$ | 0 |

which implies: $\text{Cov}(X_1, Y_1) = 0 - \frac{K_1}{N} \cdot \frac{K_2}{N}$, and that of

| $X_1=$ / $Y_2=$ | 0 | 1 |
|---|---|---|
| 0 | rest | $\frac{K_1}{N} \cdot \frac{N-1-K_2}{N-1}$ |
| 1 | $\frac{K_2}{N} \cdot \frac{N-1-K_1}{N-1}$ | $\frac{K_1}{N} \cdot \frac{K_2}{N-1}$ |

which similarly yields: $\text{Cov}(X_1, Y_2) = \frac{K_1 K_2}{N(N-1)} - \frac{K_1}{N} \cdot \frac{K_2}{N} = \frac{K_1 K_2}{N^2(N-1)}$. Putting it together: $\text{Cov}(X,Y) = -n\frac{K_1 K_2}{N^2} + n(n-1)\frac{K_1 K_2}{N^2(N-1)} = \boxed{-n\frac{K_1}{N}\frac{K_2}{N}\frac{N-n}{N-1}}$. Note that this is like the multinomial covariance of $-np_x\,p_y$, further multiplied by the correction factor of the old $\text{Var}(X)$ formula.

One can easily extend this covariance formula to the important 'overlapping' case of $U$ and $V$:

$$\text{Cov}(U,V) = -n\left(\frac{K_1}{N}\frac{K_2}{N} - \frac{K_{12}}{N}\right)\frac{N-n}{N-1}$$

where $K_1$ and $K_2$ is the number of objects contributing to $U$ and $V$ respectively, and $K_{12}$ is the number contributing to *both*.

EXAMPLES:

- Pay \$15 to play the following game: 5 cards are dealt from the ordinary deck of 52 and you get paid \$20 for each ace, \$10 for each king and \$5 for each queen. Find the expected value of your net win, and its standard deviation. We introduce $X$, $Y$ and $Z$ for the number of aces, kings and queens dealt. The net win is: $W = 20X + 10Y + 5Z - 15$, which has the expected value of $20\mathbb{E}(X) + 10\mathbb{E}(Y) + 5\mathbb{E}(Z) - 15 = 20 \times 5 \times \frac{4}{52} + 10 \times 5 \times \frac{4}{52} + 5 \times 5 \times \frac{4}{52} - 15 = -1.538$ and variance equal to $20^2\text{Var}(X) + 10^2\text{Var}(Y) + 5^2\text{Var}(Z) + 2\times 20\times 10\times\text{Cov}(X,Y) + 2\times 20\times 5\times\text{Cov}(X,Z) + 2\times 10\times 5\times\text{Cov}(Y,Z) = [5\times(400+100+25)\times\frac{1}{13}\times\frac{12}{13} - 5\times(400+200+100)\times\frac{1}{13}\times\frac{1}{13}]\times\frac{47}{51} = 152.69$ Answer: $-1.538 \pm 12.357$ dollars.

- Five cards are dealt, and we get \$1 for each spade and \$2 for each diamond, but we have to pay \$10 for each ace. Find the expected value and standard deviation of the net win. Introduce $X$, $Y$ and $U$ for the number of spades, diamonds and aces. $W = X + 2Y - 10U$, the corresponding expected value is $5 \times \frac{13}{52} + 2 \times 5 \times \frac{13}{52} - 10 \times 5 \times \frac{4}{52} = -0.09614$ and $\text{Var}(W) = 5 \times [\frac{1}{4}\times\frac{3}{4} + 2^2 \times \frac{1}{4}\times\frac{3}{4} + (-10)^2 \times \frac{1}{13}\times\frac{12}{13} - 2\times 2\times\frac{1}{4}\times\frac{1}{4}] \times \frac{47}{51} = 35.886$. Answer: $-0.0961 \pm 5.9905$ dollars

'Mixed' EXAMPLES:

- A die is rolled 5 times and we are paid \$2 for each dot obtained, then a coin is flipped 10 times and we have to pay \$7 for each head shown. Find the expected value and standard deviation of the game's net win. $\triangleright$ Let $X_1$, $X_2, ....X_5$ represent the number of dots obtained in each roll, and $Y$ be the total number of heads shown. Then $W = 2(X_1 + X_2 + ... + X_5) - 7Y$, having the expected value of $2(3.5 + 3.5 + .... + 3.5) + 7 \times 10 \times \frac{1}{2} = 0$ (a fair game) and $\text{Var}(W) = 2^2 \times (\frac{35}{12} + \frac{35}{12} + ... + \frac{35}{12}) + 7^2 \times 10 \times \frac{1}{2} \times \frac{1}{2} = 180.8\bar{3}$. Answer: $0 \pm 13.45$ dollars.

- Pay \$35, then roll a die until the $3^{rd}$ six is obtained and be paid \$2 for each roll. Find $\mu_w$ and $\sigma_w$, where $W$ is your net win. $\triangleright$ This time we introduce only $X$: the number of rolls to get the $3^{rd}$ six. Obviously $W = 2X - 35$,with the mean of $2 \times \frac{3}{\frac{1}{6}} - 35 = 1$ and $\text{Var}(W) = 2^2 \times 3 \times 6 \times 5$. Answer: $1 \pm 18.97$ dollars. Supplementary: What is the expected value and standard deviation of the net win after *15 rounds* of this game? $\triangleright$ The games are obviously played independently of each other, therefore $\mathbb{E}(W_1 + W_2 + .... + W_{15}) = 15\mu_w$ and $\text{Var}(W_1 + W_2 + ... + W_{15}) = 15\text{Var}(W)$. Answer: $15 \pm 73.485$ dollars.