Commonly Tested Discrete Distributions

Standard Discrete Distributions – We will use the shortcut notation

$$f_X(x) = \Pr(X = x) = p_x .$$

1. $X \sim U(N)$ (Uniform Distribution on the set $\{1,2,\ldots,N\}$.)
   X = the outcome of an experiment in which each outcome is equally likely

$$Supp(X) = \{1,2,\cdots,N\}$$
$$p_x = \frac{1}{N}$$
$$E[X] = \frac{N+1}{2}$$
$$Var(X) = \frac{N^2-1}{12}$$
$$M_X(t) = \sum_{j=1}^{N} \frac{e^{jt}}{N} = \frac{e^t(e^{N\cdot t}-1)}{N(e^t-1)}$$

2. $X \sim B(n,p)$ (Binomial Distribution with parameters n and p.)
   p = Pr(success) in each individual trial
   X = the number of successes in n independent trials

$$Supp(X) = \{0,1,2,\cdots,n\}$$
$$p_x = C(n,x)\cdot p^x \cdot q^{n-x} = \frac{n!}{x!\cdot(n-x)!}\cdot p^x \cdot (1-p)^{n-x}$$
$$E[X] = n\cdot p$$
$$Var(X) = n\cdot p\cdot q = n\cdot p\cdot(1-p)$$
$$M_X(t) = (1-p+p\cdot e^t)^n$$

The special case where $n = 1$ is called the Bernoulli Distribution.

3. $X \sim G(p)$ (Geometric Distribution with parameter *p*.)
p = Pr(success) in each individual trial
X = the number of failures before the first success

$$Supp(X) = \{0,1,2,\cdots\}$$

$$p_x = q^x \cdot p = (1-p)^x \cdot p$$

$$E[X] = \frac{q}{p} = \frac{1-p}{p}$$

$$Var(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$$

$$M_X(t) = \frac{p}{1-q\cdot e^t} = \frac{p}{1-(1-p)\cdot e^t}$$

4. $X \sim NB(r, p)$ (Negative Binomial Distribution with parameters *r* and *p*.)
p = Pr(success) in each individual trial
X = the number of failures before the r^{th} success

$$Supp(X) = \{0,1,2,\cdots\}$$

$$p_x = C(x+r-1, r-1)\cdot p^r \cdot q^x = \frac{(x+r-1)!}{x!\cdot(r-1)!}\cdot p^r \cdot (1-p)^x$$

$$E[X] = \frac{r\cdot q}{p} = \frac{r\cdot(1-p)}{p}$$

$$Var(X) = \frac{r\cdot q}{p^2} = \frac{r\cdot(1-p)}{p^2}$$

$$M_X(t) = \left[\frac{p}{1-q\cdot e^t}\right]^r = \left[\frac{p}{1-(1-p)\cdot e^t}\right]^r$$

This distribution is a generalization of the geometric distribution. That is, the negative binomial distribution with *r* = 1 is the same as the geometric distribution. In symbols,

$$NB(1, p) = G(p).$$

5. $X \sim P(\lambda)$ (Poisson Distribution with parameter λ.)

X = the number of occurrences of an event in a certain time period

$$Supp(X) = \{0,1,2,\cdots\}$$
$$p_x = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$
$$E[X] = \lambda$$
$$Var(X) = \lambda$$
$$M_X(t) = e^{\lambda(e^t - 1)}$$

Unlike problems involving the other distributions where you are typically given a description of the random variable and you must determine from the description which random variable you are dealing with, in problems involving the Poisson distribution you will be told that the random variable follows a Poisson distribution.

If the occurrences of the event are independent (which is typically the case), then changing the time period equates to changing the parameter.

6. $X \sim HG(M, K, n)$ (Hypergeometric Distribution with parameters *M*, *K*, and *n*.)

M = the total number of objects under consideration
K = the number of objects of certain type 1 (so M-K are not of type 1)
n = the number of objects that will be pulled from the M total objects
X = the number of the n objects that are type 1

$$Supp(X) = \{\max(0, n-(M-K)), \max(0, n-(M-K))+1, \cdots, \min(K,n)\}$$
$$p_x = \frac{C(K,x) \cdot C(M-K, n-x)}{C(M,n)}$$
$$E[X] = \frac{n \cdot K}{M}$$
$$Var(X) = \frac{n \cdot K \cdot (M-K) \cdot (M-n)}{M^2 \cdot (M-1)} = E[X] \cdot \frac{(M-K) \cdot (M-n)}{M \cdot (M-1)}$$

7. $X_i \sim M(n, p_1, \cdots, p_k)$ (Multinomial distribution with parameters n, p_1, … , p_k.)
n = number of (independent) times an experiment is performed
k = the number of possible outcomes each time experiment is performed
p_j = Pr(outcome of experiment = j), where $j = 1,2,\ldots,k$
X_i = the number of experiments that result in outcome i

$$p(x_1, x_2, \cdots, x_k) = \Pr(X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k) = \frac{n!}{x_1! \cdot x_2! \cdots x_k!} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k}$$

$$E[X_i] = n \cdot p_i$$

$$Var(X_i) = n \cdot p_i \cdot q_i = n \cdot p_i \cdot (1 - p_i)$$

$$Cov(X_i, X_j) = n \cdot p_i \cdot p_j$$

Tweaks To Some Of These Distributions

Uniform Distribution: If the support of the distribution is a set other than $\{1,2,\ldots,N\}$, then we need to tweak the formulas above.

You will find the following formula useful in this context.

$$1^2 + 2^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}$$

Geometric Distribution: The standard geometric random variable has X = the number of failures until the first success. Alternatively, we could define the random variable Y = the number of the trial in which the first success occurs. Then Y also has a geometric distribution but we need to tweak the formulas. Note that $Y = X + 1$, and so

$$p_y = \Pr(Y = y) = \Pr(X + 1 = y) = \Pr(X = y - 1) = p_{y-1} = q^{y-1} \cdot p = (1 - p)^{y-1} \cdot p$$

$$E[Y] = E[X + 1] = E[X] + 1 = \frac{q}{p} + 1 = \frac{1 - p}{p} + 1 = \frac{1}{p}$$

$$Var(Y) = Var(X + 1) = Var(X) = \frac{q}{p^2} = \frac{1 - p}{p^2}$$

$$M_Y(t) = E[e^{t \cdot Y}] = E[e^{t \cdot (X+1)}] = E[e^{t \cdot X} \cdot e^t] = e^t \cdot E[e^{t \cdot X}] = e^t \cdot M_X(t) = \frac{p \cdot e^t}{1 - q \cdot e^t} = \frac{p \cdot e^t}{1 - (1 - p) \cdot e^t}$$

Negative Binomial Distribution: Since the negative binomial distribution is a generalization of the geometric distribution, there are formulas similar to the above formulas for the negative binomial distribution. Instead of working through more formulas, I suggest interpreting all negative binomial questions in the standard form described earlier in this section.