(Short) Solutions to Problem Set #9.

(1) Let the random variable Y be the number of hypertensives in one family. If we assume the hypertensive status of siblings are independent, Y is binomially distributed:

$$P(Y = k) = \binom{n}{k} p^k q^{n-k}$$

with $n = 3$ (the number of siblings) and $p = 0.18$ (the probability of each sibling being hypertensive) (as always, $q = 1 - p = 0.82$). Hence, $P(Y = 0) = (0.82)^3 = 0.551368$, $P(Y = 1) = 3(0.82)^2(0.18)^1 = 0.363096$, $P(Y = 2) = 3(0.82)^1(0.18)^2 = 0.079704$, and $P(Y = 3) = (0.18)^3 = 0.005832$.

(2) Note that for one sibship, the probability of having at least two affected siblings is $P(Y \geq 2) = P(Y = 2) + P(Y = 3) = 0.085536 \approx 0.09$.

(a) For X, we have a sample of 25 (independent) siblings, so X will also have a binomial distribution, but this time $n = 25$ (the size of the sample) and $p = P(Y \geq 2) \approx 0.09$. Since X is binomially distributed, we can use the formulas for expected value and standard deviation from Chapter IV: $\mu = E(X) = np \approx 25 \times 0.09 \approx 2.5$, and variance is $\sigma = \sqrt{npq} \approx \sqrt{25 \times 0.09 \times 0.9} \approx 1.42$.

(b) By the "95% rule", the value of X will lie in the interval $\mu - 2\sigma \leq X \leq \mu + 2\sigma$ about 95% of all time. Since $X = 5$ falls within this interval (although almost on the edge of it), probably these data are consistent.

(3) (b) The cumulative distribution is $\Phi(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$. In our case, $\Phi(x) = 0$ for $x < 0$, and $\Phi(x) = 1$ for $x > 2$. For $0 \leq x \leq 2$, $\Phi(x) = \int_0^x (t/2)dt = t^2/4$.

(c) $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)dx = \int_0^2 (x^2/2)dx = 2^3/6 = 4/3$.

(d) $Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx = \int_0^2 (x - 4/3)^2 (x/2)dx = (4/9)2^2 - (4/9)2^3 + (1/8)2^4 = 16/9 - 32/9 + 16/8 = 2 - 16/9 = 2/9$.

(4) (a) X is binomially distributed with expected value $\mu = 0$ and standard deviation $\sigma = 2$. Hence the probability density of X is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) = \frac{1}{2\sqrt{2\pi}} \exp(-\frac{x^2}{8})$$

.

So $P(|X| < 3) = \int_{-3}^{3} f(x)dx = 2\int_0^3 f(x)dx$. If we make a change of variable $t = x/2$, the integral $\int_0^3 f(x)dx$ is transformed into the one given to us, so $\int_0^3 f(x)dx \approx 0.43$ and $P(|X| < 3) \approx 0.86$.

(b) If the target is a 6×6 square, the event of hitting the target is $A = \{-3 < X < 3, -3 < Y < 3\} = \{-3 < X < 3\} \cap \{-3 < Y < 3\}$. We assume X and Y are independent, so $P(A) = P(-3 < X < 3) \times P(-3 < Y < 3)$. We already know $P(-3 < X < 3) \approx 0.86$. Since Y has the same distribution, $P(-3 < Y < 3) \approx 0.86$ and $P(A) \approx (0.86)^2 \approx 0.74$.

(c) Answer: For any target R, the probability of hitting this target is $\int\int_R f(x, y)dxdy$, where $f(x, y) = f(x)f(y)$ is kind of probability density of the vector-valued random variable (X, Y). Here $f(x)$ is given by the above formulas. For a circular taget, the integral is easy to compute in polar coordinates.

(5)(a) Using the fomulas for normal distribution (see solution for (4)(a)), we get $\mu = 1000$ and $\sigma = 4$.

(b) Let Z be the sum of weight of people in the sample, so that $Y = Z/25$. Note that Z is the sum of 25 independent variables, each with the same distribution as X. Then Chapter 5.6 tells us that Z has normal distribution with $E(Z) = 25E(X)$, $Var(Z) = 25Var(X)$. Now, $Y = Z/25$, so $E(Y) = E(Z)/25 = E(X)$ and $Var(Y) = Var(Z)/25^2 = Var(X)/25$. Hence $E(X) = E(Y) = 1000$, $Var(X) = 25Var(Y) = 25\sigma^2 = 400$, and X has expected value 1000 and standard deviation $\sqrt{400} = 20$ (obviously, there's a typo in this problem).