Solutions to the Random Collection of Problems

Math 394

Quizzes

1

Suppose A and B are such that $P[A] = .6, P[B] = .8, P[A|B] = .375$. Then, $P[A \cup B] =$

○ .6

○ .8

○ .95

⊗ None of the above - the system is not coherent . You will notice that $P[A \cap B] = P[B][A|B] = .8 * .375 = .3$, so that $[A \cup B] = P[A] + P[B] - P[A \cup B] = 1.4 - .3 = 1.1 > 1!$

2

Consider n independent Bernoulli RVs, all with parameter p. The distribution of the RV $Y = X_1 \cdot X_2 \cdot \ldots \cdot X_n$ is

○ Binomial $bin(n, p)$

○ The same as the distribution of e^{N_n}, where N_n is a binomial $bin(n, p)$

○ $P[Y = k] = p^k \quad k = 0, 1, \ldots, n$

○ $P[Y = k] = \frac{p^k}{(1-p)^{1-k}} \quad k = 0, 1$

⊗ $P[Y = k] = p^n \; k = 1, \quad P[Y = k] = 1 - p^n \; k = 0$. A product of 0s and 1s can only be 0 or 1 - but will not be 1 unless all factors are equal to1.

○ $P[Y = k] = p^n \; k = 1, \quad P[Y = k] = (1 - p)^n \; k = 0$

3

Recalling the definition of variance for a RV, what is the *general* formula for $Var[X-Y]$, where X, and Y are two general RVs (no special assumptions are made, except that their first and second moments are all well defined). You have your book, so you shouldn't need this, but, anyway, here goes:

$$Var[X] = E\left[X^2\right] - (E[X])^2$$

$$Cov[X,Y] = E[XY] - E[X]\,E[Y]$$

○ $Var[X] + Var[Y]$

○ $Var[X] - Var[Y]$

○ $Var[X] + Var[X] + Cov[X,Y]$

○ $Var[X] + Var[X] - Cov[X,Y]$

○ $Var[X] + Var[X] + 2Cov[X,Y]$

⊗ $Var[X]+Var[X]-2Cov[X,Y]$. It's possibly faster to do with the alternate definition (there is no substantial difference, anyway), and

$$Var[X-Y] = E\left[(X - Y - E[X-Y])^2\right] =$$

$$E\left[(X - E[X])^2 + (Y - E[Y])^2 - 2(X - E[X])(Y - E[Y])\right] =$$

$$= Var[X] + Var[X] - 2Cov[X,Y]$$

4

If A and B are independent, then...

Solution: the only possible correct answers would be

$$P[A \cap B] = P[A]\,P[B]$$

or

$$P[A|B] = P[A]$$

Beware of answers like $P[A \cap B] = 0$ which are especially false... (in the sense that it is a serious and frequent mistake)

If $A \subseteq B$, then ($P[A|B] =$, $P[B|A] =$, $P[A \cup B] =$, $P[A \cap B] =$...)

Solution: If $A \subseteq B$, then $P[A \cap B] = A$, and $P[A \cup B] = B$. Hence, for instance,

$$P[B|A] = \frac{P[A \cap B]}{P[A]} = \frac{P[A]}{P[A]} = 1$$

(you see, $\subseteq$is connected to the logic "implies", and if an outcome is in A, it certainly will be in B, so A implies B), while

$$P[A|B] = \frac{P[A]}{P[B]}$$

(which is also pretty intuitive)

5

Suppose $X_1, X_2, \ldots, X_n$ are independent Bernoulli RVs. Then the distribution of

$X_1^2 + X_2^2 + \ldots + X_n^2$ is...
$e^{X_1} + e^{X_2} + \ldots + e^{X_n}$ is...
$e^{X_1+X_2+\ldots+X_n}$ is...

Solutions: The trick is to see what values are taken, and with which probability, by the functions we are presented with. For instance, if $X_i = 0$ or 1, $X_i^2 = X_i$, so that the first sum has a binomial distribution $bin(n,p)$, just like the sum of the X_i. Note that, if the X_i take the values ± 1, then all $X_i^2 = 1$, and the sum would have been equal to n with probability one. In the case of the sum of the e^{X_i}, the summands take values e if $X_i = 1$, and 1 if $X_i = 0$. Hence,

$$P\left[e^{X_1} + e^{X_2} + \ldots + e^{X_n} = ke + (n-k)\right] = \binom{n}{k} p^k (1-p)^{n-k}$$

Finally, we have that $e^{X_1+X_2+\ldots+X_n}$ has the same distribution as e^Y, where $Y \sim bin(n,p)$. Thus,

$$P\left[e^{X_1+X_2+\ldots+X_n} = e^k\right] = \binom{n}{k} p^k (1-p)^{n-k}$$

6

Let X be a standard normal RV. Then
$P[|X| < 1] =$
$P[|X| < 2] =$
$P[|X| > 2] =$...

Solution: These are simple exercises in reading tables, or punching your calculator. So

$$p(x) = P[|X| < x] = 1 - P[|X| > x] = 2\Phi(x) - 1$$

and we have for

- $x = 1$ $\quad p(x) = .68269$
- $x = 2$ $\quad p(x) = .95450$
- $x = 3$ $\quad p(x) = .99730$, etc.

7

If A and B are independent events, then (there could be more than one correct answer)

○ $P[A \cup B] = P[A] + P[B]$

⊗ $P[A \cap B] = P[A]\,P[B]$It's just the definition.

○ $P[A \setminus B] = P[A] - P[B]$

○ $P[A] = 0$, or $P[B] = 0$, or both Note that if any of the conditions mentioned are true, A and B would be independent (in the sense that their intersection would have 0 probability, equal to the product of zero with anything else), but you obviously would not infer that one of the sets has zero probability, just because you knew they were independent!

8

If X is a RV with binomial distribution with parameters (n, p), then $P[nX = k] =$(there could be more than one correct answer)

○ $\binom{n^2}{k} p^k (1-p)^{n-k}$

○ $\binom{n}{nk} p^{nk} (1-p)^{n-nk}$

⊗ $\binom{n}{\frac{k}{n}} p^{\frac{k}{n}} (1-p)^{n-\frac{k}{n}}$Since $nX = k \Rightarrow X = \frac{k}{n}$, this follows immediately. As usual, we'll read the formula to mean 0 if $\frac{k}{n}$ is not an integer.

○ $\binom{\frac{n}{k}}{k} p^k (1-p)^{\frac{n}{k}-k}$

○ $\binom{n^2}{nk} p^{nk} (1-p)^{n^2-nk}$

○ $\binom{n}{k} p^k (1-p)^{n-k}$

9

Suppose X is a normally distributed RV, with $\mu = 1, \sigma^2 = 2$, then (there could be more than one correct answer)

⊗ $P[X < 1] = P[X > 1]$The normal distribution is symmetric around its expected value: both probabilities here are equal to $\frac{1}{2}$.

○ $P[X < 0] = P[X > 0]$This and the following answer would both be true if $\mu = 0$, but it is not.

○ $P[X < -2] = P[X > 2]$

○ $P[X < 2] = P[X > 2]$This would be true if $\mu = 2$. Note that the value of σ^2 is not relevant for these questions.

Problems

1

A test for defects in a shipment is performed on a sample, and this procedure is known, from statistical theory, to have a probability of 5% of wrongly rejecting the shipment, and a probability of 10% of wrongly accepting the shipment. The original manufacturer is known to have a 2% rate of defective shipments.

1. When the test result is negative (hence, the shipment is accepted), what is the probability that the shipment is actually defective?
2. When the test result is positive (hence, the shipment is rejected), what is the probability that the shipment is indeed defective?
3. If the procedure is amended to include a second test (independent, but with the same correctness as the first one) when the test is failed, and the shipment is rejected only if both tests are positive, what are the probabilities that a good shipment will be rejected, and that a bad shipment will be accepted?

Solution: Let $X = 1$ if the shipment is good, and 0 if it is not. Let $Y = 1$ if the test is negative (shipment looks OK), and 0 if it is not. The problem says that

$$P[X = 0] = .02$$

$$P[Y = 1 | X = 1] = .95; P[Y = 1 | X = 0] = .1$$

1. We are asked for $P[X = 0 | Y = 1]$. This is

$$= \frac{P[Y = 1 | X = 0] P[X = 0]}{P[Y = 1]} = \frac{P[Y = 1 | X = 0] P[X = 0]}{P[Y = 1 | X = 1] P[X = 1] + P[Y = 1 | X = 0] P[X = 0]} =$$

$$= \frac{.1 \cdot .02}{.95 \cdot .98 + .1 \cdot .02} = .0021436 \qquad (1)$$

or about .2%.

2. Now it is $P[X = 0 | Y = 0]$. Similarly,

$$= \frac{P[Y = 0 | X = 0] P[X = 0]}{P[Y = 0]} = \frac{P[Y = 0 | X = 0] P[X = 0]}{P[Y = 0 | X = 1] P[X = 1] + P[Y = 0 | X = 0] P[X = 0]} =$$

$$\frac{.9 \cdot .02}{.05 \cdot .98 + .9 \cdot .02} = .26866$$

3. Now, if we think in terms of a tree, we have to compute the probability that a good shipment is rejected, and a bad is accepted. Starting with the good shipment, it will be rejected, if two tests fail. Hence we are looking at

$$(P[Y = 0 | X = 1])^2 = .5^2 = .25$$

Note that the point here is that we *know the shipment is good*. Similarly, it is easier to compute the probability of rejecting a bad shipment

$$(P[Y = 0 | X = 0])^2 = .9^2 = .81$$

and the probability of accepting it will be $1 - .81 = .19$

Note You might be thinking of a different question, such as "what is the probability that a rejected shipment is bad". Since we just found what the conditional probabilities of a double failed test are, we can easily answer that question too:

$$P[X=0\,|Y_1=1,Y_2=0]=\frac{P[\text{reject}\,|X=0]\,P[X=0]}{P[\text{reject}]}=$$

$$\frac{.81\cdot .02}{.81\cdot .02+.25\cdot .98}=.93388$$

more of them getting by. In fact, a passed shipment will be good with probability

$$P[X=1\,|\{Y_1=1\}\cup\{Y_1=0,Y_2=1\}]=$$

$$=\frac{P[\{Y_1=1\}\cup\{Y_1=0,Y_2=1\}\,|X=1]\,P[X=1]}{P[\{Y_1=1\}\cup\{Y_1=0,Y_2=1\}]}=$$

$$\frac{.95\cdot .98+.95\cdot .05\cdot .98}{.95\cdot .98+.95\cdot .05\cdot .98+.02\cdot .1+.02\cdot .9\cdot .1}=.99613$$

still very good, but not as good as $1-.0021436=.99786$ (from (1))

2

Suppose X_j $j=1,2,\ldots$ are RVs with distribution

$$P[X_j=1]=p;\,P[X_j=-1]=1-p$$

and that they are independent. Let $S_n=\sum_{k=1}^{n}X_k$

1. Write the pmf of S_n
2. Write $E[S_n]\,,Var[S_n]$, and the approximate distribution (e.g., the cdf, or something equivalent) for $\frac{S_n-ES_n}{\sqrt{Var[S_n]}}$ when n is very large.
3. Assuming now that the parameter p is of the form $p_n=\frac{p}{n}$, for some number p, write the approximate pmf for S_n, when n is very large. *Note: be very careful here! You need to adjust what you can find in the book, in order to fit the present situation.*
4. Suppose $n=1000$, $p=.3$. Use one of the results above to evaluate explicitly $P[S_n\leq 500]$

Solution: We are considering a sum of i.i.d. "Bernoulli" RVs, whose values are not 1 and 0, but 1 and -1 (like "win/loss" or "spin", rather than "count the wins).

1. $S_n=W_n-L_n$, where W_n is the number of "+1" in n trials, and L_n the number of "-1". Hence $L_n=n-W_n$, while

$$P[W_n=k]=\binom{n}{k}p^k(1-p)^{n-k}$$

and

$$P[S_n=j]=P\left[W_n=\frac{n+j}{2}\right]\qquad(2)$$

(the probability being zero if $\frac{n+j}{2}$ is not an integer). The pmf is thus

$$P[S_n=j]=\binom{n}{\frac{n+j}{2}}p^{\frac{n+j}{2}}(1-p)^{\frac{n-j}{2}}$$

($n-\frac{n+j}{2}=\frac{n-j}{2}$). The expression is assumed to be zero when $\frac{n\pm j}{2}$ is not an integer.

2. Clearly

$$EX_j = p - (1-p) = 2p - 1$$

$$Var[X_j] = p(1-(2p-1))^2 + (1-p)(-1-(2p-1))^2 =$$
$$= p \cdot 4p^2 + 4(1-p)\left(p^2+2p+1\right) = 4\left(1+p-p^2\right)$$

so that

$$ES_n = 2np - n; Var[S_n] = 4n\left(1+p-p^2\right)$$

The Central Limit Theorem implies that, for n very large,

$$P\left[\frac{S_n - ES_n}{\sqrt{Var[S_n]}} \le s\right] \simeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{-\frac{x^2}{2}} dx$$

3. Essentially, we are looking at a Poisson approximation. However, we must adjust the calculations, similarly to what we did in point 1: In fact, we can use a Poisson approximation for W_n, and apply then (2) to get the distribution of S_n. Indeed, W_n is the sum n standard Bernoulli RVs, with $p_n = \frac{p}{n}$, and hence, for large n, will be well approximated by a Poisson RV, with parameter p. I.e.,

$$P[W_n = k] \simeq \frac{p^k}{k!} e^{-p}$$

Applying (2)

$$P[S_n = j] = P\left[W_n = \frac{n+j}{2}\right] \simeq \frac{p^{\frac{n+j}{2}}}{\left(\frac{n+j}{2}\right)!} e^{-p}$$

where, again, we must agree to consider the expression equal to zero when $\frac{n+j}{2}$ is not an integer.

4. With these values, $ES_n = 300$, and $Var[S_n] = 4 \cdot 10^3 \cdot (1 + .3 - .09) = 4840$, so the better approximation seems to be the Gaussian one. Using the result from point 2,

$$P[S_n \le 500] = P\left[\frac{S_n - ES_n}{\sqrt{Var[S_n]}} \le \frac{500-300}{69.57}\right] = .99798$$

3

Suppose U is a uniform RV, with $P[0 \le U \le 1] = 1$. Let $g(x) = \log \frac{1}{x}$. Write the distribution of $g(U)$.

Hint: compute the cdf or the survival function of $g(U)$

Solution: Follow the hint, and compute

$$P\left[\log U^{-1} \le t\right] = P[-\log U \le t] = P\left[U \ge e^{-t}\right] = e^{-t}$$

as long as $0 < e^{-t} \le 1$, i.e. $0 \le t \le \infty$. Thus $\log U^{-1}$ is an exponential RV. This method is often used to "produce" an exponential RV when one is given a uniform RV, specifically by a computer routine. Pseudo-code would look something like

```
u=rnd();
x=-log(u);
```

4

A binary signal is transmitted, each bit having, independently, a 10% chance of being received wrongly.

1. If each packet is composed of 10 bits, what is the distribution of the number of errors in the reception?
2. What is the probability that a packet arrives corrupted (i.e., not all bits are received correctly)?
3. Suppose the 10th bit is a "parity" bit, i.e. a 1 of the other 9 bits have an odd number of 1s, and a 0 if they have an even number (0 is even). What is the probability that a corrupted packet will go undetected?

Solution: This eventually gets a bit involved, but it starts easy enough:

1. Clearly, by independence, the distribution of the number of errors over 10 bits is $bin(10, .1)$
2. The probability of corruption is $1 - P[\text{no error}]$, i.e., if E is the number of errors,
$$1 - P[E = 0] = 1 - .9^{10} = .65132$$
3. This is a standard (minimal) "ancient" error-checking procedure. Originally, when all ASCII characters could be represented in 7 bits, some protocols would use an 8th bit as a parity bit. 8 bits are needed for the "extended ASCII" code, the final bit is sometimes used as a "stop" bit - signaling the end of the packet - and one would indeed have 10 bits, if one added a parity bit - in reality, 8-bit protocols over voice lines, usually went 8-no stop-no parity. A good packet is one in which no bits were corrupted, and the parity bit is correctly set. It is easy to see that the parity bit will erroneously tell us that all is fine if an even number of bits were reversed (including, possibly, the parity bit itself). Hence, we will have a false "all clear" if there were 2,4,6,8, or 10 errors. This happens with probability
$$\binom{10}{2}.1^2 \cdot .9^8 + \binom{10}{4}.1^4 \cdot .9^6 + \binom{10}{6}.1^6 \cdot .9^4 + \binom{10}{8}.1^8 \cdot .9^2 + \binom{10}{10}.1^{10} =$$
$$= .20501$$

5

A rat in a maze is confronted by a sequence of forks, where it has to choose between going right or left. To test its sense of smell, a piece of cheese is placed at the left end of a 20-fork maze, and its choices are recorded, in order to see if they show a left-ward bias, compared to a no-cheese maze.

The results are:

- no-cheese: 12 left turns, 8 right turns
- cheese: 14 left turns, 6 right turns

1. Assuming no bias is really present, i.e. the probabilities of a left or right turn are $p = \frac{1}{2}$, what are the probabilities that the rat would make a number of left turns, $n \geq 12$, and $n \geq 14$? Write an exact formula, and use an approximation to get an explicit number.

2. The theory of the behavioral scientists was that the difference between the two setups (left-right), would have been distributed (in the Gaussian approximation) with a $\mu = 5$, $\sigma = 2$ (note: that is σ, not σ^2). If that model was correct, what would be the probability that we would observe a difference (like we did) $D \leq 2$?

Solution: This is a straight binomial distribution for the number of left (or right) turns.

1. Hence,

$$P[L \geq 12] = \sum_{k=12}^{20} \binom{20}{k} \frac{1}{2^{20}}$$

and

$$P[L \geq 14] = \sum_{k=14}^{20} \binom{20}{k} \frac{1}{2^{20}}$$

Rather than computing these expressions, it is much more relaxing to compute

$$P\left[\frac{L-10}{\sqrt{20 \cdot \frac{1}{4}}} \geq \frac{12-10}{\sqrt{5}}\right], P\left[\frac{L-10}{\sqrt{20 \cdot \frac{1}{4}}} \geq \frac{14-10}{\sqrt{5}}\right]$$

assuming that the first fraction is a standard normal RV ($E[L] = np = 20 \cdot \frac{1}{2}$, and $Var[L] = np(1-p) = 20 \cdot \frac{1}{2} \cdot \frac{1}{2} = 5$). The result is, respectively,

$$P[L \geq 12] \simeq .18555$$

$$P[L \geq 14] \simeq .036819$$

2. Now, D would be the difference of two (assumed independent) normal RVs. We have noted in problems, and elsewhere, that sums and differences of independent normals are normal - the expected values sum or subtract, according to the operation, and the variances always sum (it is very peculiar of normal RVs that their sum is Gaussian even when they are not independent - they have, however, to exhibit what is called a *jointly Gaussian* distribution - though, while the expectations sum or subtract, as always, the formula for the variance needs to take the covariance into account as well). Since D is supposed to be $N(5, 4)$,

$$P[D \leq 2] = P\left[\frac{D-5}{2} \leq \frac{2-5}{2}\right] = \Phi(-1.5) = .066807$$

6

A simple model of reliability for an electronic equipment would be the following:

- With probability $p = .01$ the equipment is DOA, i.e. is non operative from the beginning
- Assuming it is *not* DOA, it starts with a *hazard function* of the form

  $$h_1(t) = 1 - t; 0 \le t \le 0.5$$

  then changes to

  $$h_2(t) = 0.5; 0.5 \le t \le 2$$

  and finally ends with

  $$h_3(t) = 0.5 + (t - 2); t \ge 2$$

1. Write the cdf for the lifetime of the equipment.
2. How likely is the equipment going to fail between $t = 0$ and $t = 1$?
3. Assume the component comes with a full warranty up to time $t = 1$. If the cost of replacement is $\$C = \$1,000$, what is the expected cost of this warranty for the manufacturer?
4. *Assuming the component has not failed before* $t = 1$, an extended warranty is offered for full replacement at \$1,000, for the time period $1 \le t \le 3$: what is the expected cost of this for the warranty company?

Solution: This is a little mean, since the distribution of the lifetime is "mixed". Call the lifetime of an equipment L. We are told that

$$P[L = 0] = .01$$

$$P[L > t \,|L > 0] = e^{-\int_0^t h(s)ds} (t > 0)$$

(Note the phrasing about the case the piece is not DOA: it mentions a *conditional* distribution).

1. The cdf is easily found from the survival function, which goes like this:

$$R_L(0) = P[L > 0] = .99$$

$$R_L(t) = P[L > t] = P[L > t \,|L > 0]\, P[L > 0] =$$

$$e^{-\int_0^t h(s)ds} \cdot .99$$

Hence,

$$R_L(t) = \begin{cases} t < 0 & 1 \\ t = 0 & .99 \\ 0 < t \le .5 & .99e^{-\int_0^t (1-s)ds} \\ .5 < t \le 2 & .60047e^{-\int_{.5}^t .5ds} \\ 2 < t & .28364e^{-\int_2^t (s-2)ds} \end{cases}$$

$$F_L(t) = \begin{cases} t < 0 & 0 \\ t = 0 & .01 \\ 0 \le t < .5 & 1 - .99e^{-\int_0^t (1-s)ds} \\ .5 \le t < 2 & 1 - .60047e^{-\int_{.5}^t .5ds} \\ 2 \le t & 1 - .28364e^{-\int_2^t (s-2)ds} \end{cases}$$

2. The probability of failure between 0 and 1 (excluding 0) is

$$F(1) - F(0) = 1 - .60047e^{-.25} - 0.1 = .52236$$

We add $F(0)$, if we want to include the $L = 0$ case:

$$F(1) = .52236 + .01 = .53236$$

3. The probability of failure within the warranty period is $F_L(1)$, since the DOA case is obviously included. Hence, the expected cost of the warranty is

$$C \cdot F_L(1) = 1000 \cdot .53236 = \$532.36$$

4. This warranty kicks in if the first time period has gone by without failures. Hence, we are looking at

$$P[L \le 3 \,|L > 1] = \frac{P[1 < L \le 3]}{P[L > 1]} = \frac{F_L(3) - F_L(1)}{R_L(1)} =$$

$$\frac{F_L(3) - 1}{R_L(1)} + 1 = 1 - \frac{R_L(3)}{R_L(1)} = 1 - \frac{.28364e^{-\int_2^3 (s-2)ds}}{.60047e^{-.25}} = 1 - \frac{.28364}{.60047} = .52763$$

7

Two theories are competing as explanation of a phenomenon. One assumes that when input $X = n$ (n is an integer), the output Y should be an exponential RV with parameter $\frac{1}{n}$. The other assumes that, for the same input, the output should be Gaussian with parameters $\mu = n$, $\sigma^2 = n^2$.

1. How would the two theories differ in predicting the likelihood of an output in the range $Y > 2n$, for $X = n$.

2. A blind experiment is performed and an output of $Y \in (9, 12)$. Using tables and calculator, find the conditional probability for $X = 10$, according to both models, with $P[X = n] = \frac{1}{10}; n = 1, 2, \ldots, 10$.

Solution: Note that both theories assume the same value for mean and variance.

1. For an exponential RV with parameter $\frac{1}{n}$ (i.e., expected value n)

$$P[Y > 2n] = e^{-\frac{2n}{n}} = e^{-2} = .13534$$

For a normal RV, with distribution $N(n, n^2)$,

$$P[Y > 2n] = \frac{1}{\sqrt{2\pi \cdot n^2}} \int_{2n}^{\infty} e^{-\frac{(x-n)^2}{2n^2}} dx$$

or

$$P\left[\frac{Y-n}{n} > \frac{2n-n}{n}\right] = P[Z > 1] = 1 - \Phi(1) = .15866$$

(here Z is a standard normal RV). The two values are different, but not by much.

2. We are looking for $P[X = 10 \,|9 < Y < 12]$ for the two models. This is, by Bayes' Rule

$$= \frac{P[9 < Y < 12 \,|X = 10]\, P[X = 10]}{P[9 < Y < 12]} = p\frac{P[9 < Y < 12 \,|X = 10]\, P[X = 10]}{P[9 < Y < 12]}$$

For the exponential model,

$$P[9 < Y < 12 \,|X = n] = e^{-\frac{9}{n}} - e^{-\frac{12}{n}}$$

$$P[9 < Y < 12] = \sum_{n=1}^{10} \frac{1}{10}\left(e^{-\frac{9}{n}} - e^{-\frac{12}{n}}\right)$$

for a result of

$$\frac{e^{-\frac{9}{10}} - e^{-\frac{12}{10}}}{\sum_{n=1}^{10}\left(e^{-\frac{9}{n}} - e^{-\frac{12}{n}}\right)} \sim .15828$$

In the Gaussian case,

$$P[9 < Y < 12 \,|X = n] = \frac{1}{\sqrt{2\pi \cdot n^2}} \int_9^{12} e^{-\frac{(x-n)^2}{2n^2}} dx$$

or

$$P\left[\frac{9-n}{n} < \frac{Y-n}{n} < \frac{12-n}{n} |X = n\right] = \Phi\left(\frac{12}{n} - 1\right) - \Phi\left(\frac{9}{n} - 1\right)$$

and

$$P[9 < Y < 12] = \sum_{n=1}^{10} \frac{1}{10}\left(\Phi\left(\frac{12}{n} - 1\right) - \Phi\left(\frac{9}{n} - 1\right)\right)$$

for a result of

$$\frac{\Phi(.2) - \Phi(-.1)}{\sum_{n=1}^{10}\left(\Phi\left(\frac{12}{n} - 1\right) - \Phi\left(\frac{9}{n} - 1\right)\right)} \simeq .1$$

In other words, neither result makes a huge case for $X = 10$ (fact is, with such a flat distribution, many values of X are likely for a given output), but, as opposed to the normal model, the exponential model builds a conditional probability on this value larger than the "a priori" probability.

8

A RV has density

$$f_X = \begin{cases} cx(1-x) & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

1. What is the value of c?
2. What is the expected value of X?
3. What is the variance of X?

Solution: This is not a very creative problem. It consists of a few simple integrations...

1. Since

$$\int_0^1 x(1-x)\,dx = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

we have $c = 6$. This distribution is called the Beta(1,1) distribution.

2. Again, it is a direct integration... Or, we can look at the graph, note that it is a downward parabola, zero at $x = 0$ and $x = 1$, so that the expected value will be at the midpoint, which is a symmetry axis for the curve, $EX = \frac{1}{2}$. If you really want to do the integration,

$$6\int_0^1 x^2(1-x)\,dx = 6\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{6}{12} = \frac{1}{2}$$

3. We might as well keep up with simple integrals, and compute

$$EX^2 = 6\int_0^1 x^3(1-x)\,dx = 6\left(\frac{1}{4} - \frac{1}{5}\right) = \frac{6}{20} = \frac{3}{10}$$

and the variance is

$$EX^2 - (EX)^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$$

9

Let X_i, $i = 1, 2, \ldots, 10$ all take values $0, \pm 1$, each with probability $\frac{1}{3}$.

1. What are the possible values for $S = \sum_{i=1}^{10} X_i$?
2. Assuming that all X_i are independent, what are the probabilities $P[S = 10]$, and $P[S = -10]$?
3. Again assuming independence, what is the probability $P[S = 9]$?

Solution: Note that we now have three possible values for X_i, and this changes all of the answers, compared to the 2-values case

1. Since 0 is a possible value, S can now range through *all* integers between -10, and 10
2. The only way we can get $S = \pm 10$, is if all the X_i's take simultaneously the value 1 or -1 respectively. Due to the independence assumption, both cases happen with probability $\left(\frac{1}{3}\right)^{10} = 1.6935 \cdot 10^{-5}$
3. In order to have a value of 9 for the sum of 10 summands, we need 9 1s and a 0 (sorry for the silly mistake in the previous version). No other combination will do. There are 10 ways to have each of these two sequences (depending on the position of the non-1 entry), and all have the same probability as any given sequence, i.e., $1.6935 \cdot 10^{-5}$ (we assumed that all digits had the same probability). Hence, the answer is

$$10 \cdot 1.6935 \cdot 10^{-5} = 0.00033870$$